*Systematic Review Article*
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

*Med J Islam Repub Iran.* 2025 (16 Jun);39.81. https://doi.org/10.47176/mjiri.39.81

# AI-Powered Clinical Decision Support Systems in Disease Diagnosis, Treatment Planning, and Prognosis: A Systematic Review

Marzieh Nojomi[1], Ebrahim Babaee[1]*, Zahra Rampisheh[1], Mahshid Roohravan Benis[1], Mahdi Soheyli[1], Nasibeh Rady Raz[2]*

Received: 12 Feb 2025 Published: 16 Jun 2025

## Abstract

**Background**: Artificial intelligence (AI) is transforming healthcare with applications that can surpass human performance in prevention, detection, and treatment. This systematic review aimed to collect and assess the impact and success of AI technologies across various healthcare domains.

**Methods**: A systematic search of major databases (including PubMed, Scopus, and ISI) was conducted for articles published up to 2023. Keywords related to AI-driven disease detection, classification, and prognosis were used. Non-English articles or those with inaccessible full texts were excluded. Data was extracted by two researchers, and the quality of selected articles was evaluated based on the strengths and limitations stated by the authors.

**Results**: In total, 123 articles were included. AI contributions were categorized into three areas. For disease detection (n=75), Coronavirus disease 2019 (COVID-19) was the most frequent topic (n=18), followed by oncology. Chest X-rays were the most common input (n=15). In disease classification (n=23), oncology (especially breast cancer) was the most researched field (n=7), primarily using breast imaging. For prediction and prevention (n=25), oncology was again the most studied category, with clinical and laboratory parameters being the most utilized input (n=12).

**Conclusion**: AI-driven clinical decision support systems (CDSS) exhibit strong diagnostic and prognostic accuracy in imaging and laboratory settings. However, many models function as "black boxes," which limits interpretability and clinician trust. Data bias and challenges in integrating AI tools into practice also persist. The findings suggest that future work should focus on explainable AI and rigorous real-world validation to safely implement these tools in healthcare.

**Keywords:** Artificial Intelligence, Clinical Decision Support Systems, Diagnosis, Treatment, Prognosis, Disease, Prediction

***Conflicts of Interest:*** *None declared*
***Funding:*** *None*



***Cite this article as***: xxxxx. AI-Powered Clinical Decision Support Systems in Disease Diagnosis, Treatment Planning, and Prognosis: A Systematic Review. *Med J Islam Repub Iran.* 2025 (16 Jun);39:81. https://doi.org/10.47176/mjiri.39.81

## Introduction

Artificial Intelligence (AI) is a multidisciplinary study that aims to create a machine capable of perceiving data, inferring information, reaching intelligence, wisdom, cognition, and ultimately making decisions. AI aims to assist human beings in their decision-making by preparing a framework for processing all data at the same time, and presenting logical thinking and problem-solving (1). These machines are designed to handle vast, complex tasks without limitations in time or accuracy (2).

According to the literature, AI is applied in a broad range of applications, technologies, and facilities, from software programs to robots. Like the other services and jobs, using AI in medical care is inevitable today (3). AI could compensate for some care delivery deficiencies, such as a lack of manpower and time-consuming tasks (2). Generally, AI is used in prevention, early detection of disease, and personalized and targeted therapy. Using a variety of computational tools, AI can process different types of data, such

*Corresponding authors:* *Dr Ebrahim Babaee, babaee.e@iums.ac.ir*
*Dr Nasibeh Rady Raz, radyraz.n@iums.ac.ir*

1. Preventive Medicine and Public Health Research Center, Psychosocial Health Research Institute, Department of Community and Family Medicine, School of Medicine, Iran University of Medical Sciences
2. Department of Artificial Intelligence in Medicine, Faculty of Advanced Technologies in Medicine, Iran University of Medical Sciences, Tehran, Iran

***↑What is "already known" in this topic:***
Artificial intelligence (AI)-powered clinical decision support systems (CDSS) demonstrate high diagnostic and prognostic accuracy in imaging and laboratory settings. However, challenges such as the "black-box" nature of models, dataset biases, and integration barriers into clinical workflows limit their widespread adoption and trust among healthcare providers.

***→What this article adds:***
This systematic review provides a focused analysis of AI-powered CDSS across diagnosis, treatment, and prognosis, addressing architectural suitability (eg, CNNs for imaging, RNNs for temporal data). It critically examines limitations, proposes strategies for explainability and bias mitigation, and highlights ethical and workflow considerations for real-world deployment.

as laboratory, clinical, images, signals, et cetera, to find irregular patterns of disease, perform estimation, and prediction. The ability to work with multimodal data creates a holistic view for physicians to make quick and precise decisions.

Studies show that AI may enable better disease prevention, diagnosis, and treatment. Among the main fields of disease that use AI tools, we can mention basic interventions in the fields of cancer, neurology, cardiology, and diabetes (4-6). For instance, many studies in radiologic diagnosis of different types of lung disease find the remarkable diagnostic value of various types of AI methods (7-9). However, there is a need for research to comprehensive study and address the effect of AI applications on healthcare. Therefore, here, we systematically reviewed the evidence to find the effect of different AI methods usage on the medical interventions classified as prevention, diagnosis, and treatment.

## Methods

### *Study Design*

***Search Strategy and Study Screening Process:*** This systematic review was conducted in 2024. Iran University of Medical Sciences approved the study (IR.IUMS.REC.1401.711). To select appropriate and relevant studies, an extensive electronic search was conducted. For this purpose, PubMed, ISI, Cochrane, Scopus, Embase, Science Direct, and Elsevier databases were searched. Articles published until 2023 were searched. To select studies, by applying Mesh term strategy we used keywords such as “AI powered,” “AI powered,” “AI- powered,” “AI assisted,” “AI assisted,” “AI based,” “AI based,” “AI enabled,” “AI enabled,” “AI aided,” “AI aided,” “Machine learning powered,” “Machine learning assisted,” “Machine learning based,” “Machine learning enabled,” “Machine learning aided,” “Deep learning powered,” “Deep learning assisted,” “Deep learning based,” “Deep learning enabled,” “Deep learning aided,” “Neural Network Computer,” “Computer based,” “Computer assisted,” “Computer enabled,” “Computer aided,” “Computer powered,” “System based,” “System assisted,” “System enabled,” “System aided,” “System powered,” “AI,” “Deep learning,” “Machine learning,” "Clinical Decision Support Systems," and "Clinical Decision Support”. Based on the main purpose of the study, the keywords of prevention, classification, and diagnosis (detection) were also used to search for studies. The types of included studies were intervention clinical trials, randomized controlled trials, case-control, prospective and retrospective cohort, and cross-sectional. Also, references to the selected articles were searched manually. For an extensive search, 3 researchers conducted the resource search process separately and eventually coordinated the selected studies.

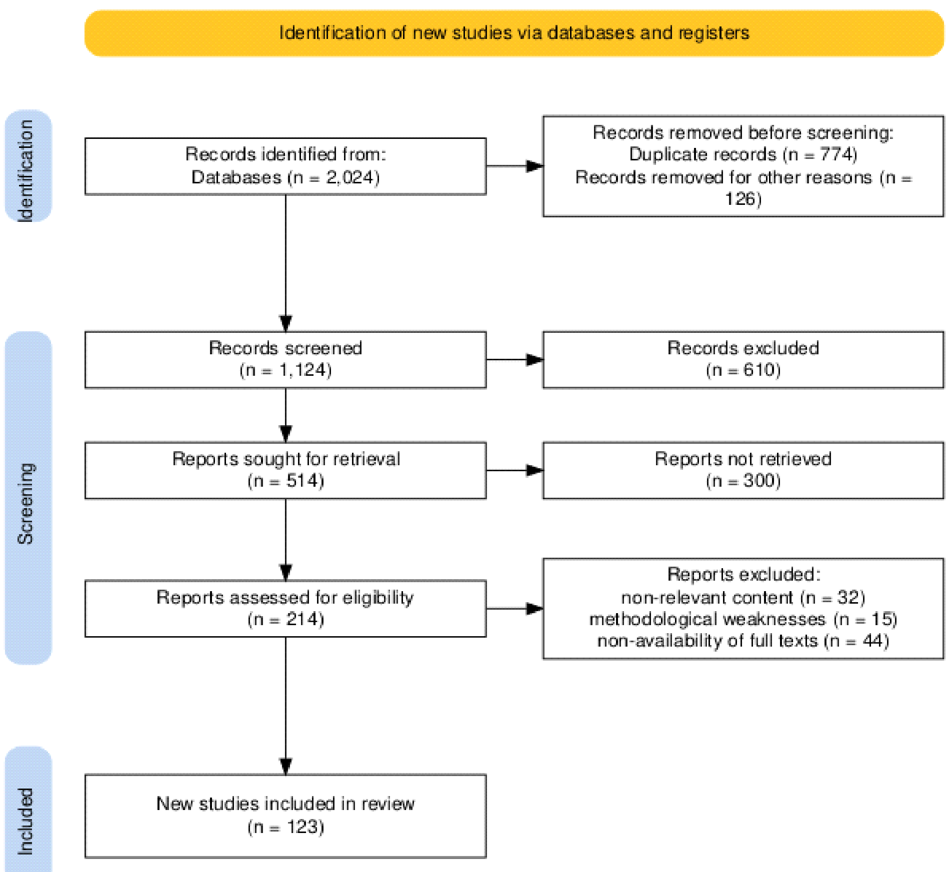


Figure 1. Finding and screening flowchart

In the first search phase, 2024 articles were selected. Duplicated articles were detected by 1 researcher and supervised by a subsequent researcher using EndNote (X17) software, and 1250 articles were removed. The criteria used for duplication detection were similar in the titles, first author name, and the year of publication. Here, we focus on the open-access journals and those publicly available for the possibility of further investigation. A total of 126 articles were excluded due to full-text unavailability. The number of remaining articles after this process reached 648. Subsequently, the titles and abstracts of articles were evaluated based on inclusion criteria, and 243 articles met the inclusion criteria. Using a full-text review, 120 articles were excluded due to inappropriate content. Finally, 123 eligible studies were reviewed. The finding and screening flowchart was plotted using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram tool (10) and reported in Figure 1.

### Inclusion and Exclusion Criteria

We considered all studies with any study design as eligible for inclusion if they examined PICO as a tool (Table 1) for developing a search strategy for identifying potentially relevant studies. We applied other restrictions in this review, such as studies related to the English language. Also, for quality assessment, we used evaluation criteria that are usual in artificial intelligence in medicine journals, including accuracy, precision, et cetera. We evaluate AI-based systems rather than the process of diagnosis. The articles whose full texts were not accessible were excluded.

### Data Extraction

A standard checklist was developed for data collection and extraction. The design of the checklist was done under the supervision of the AI expert of the project. The designed checklist included the information of the extracted articles, such as the name of authors, year of publication, the country where the study took place, the type of AI model, the type of disease, the type of data processed, the AI performance measures and the limitations that mentioned in the manuscripts. To evaluate the quality of the selected articles, the strengths and limitations expressed in the articles by the authors were used as a proxy for the quality evaluation tool of the selected studies. Data extraction was done by 2 researchers of the study under the supervision of an AI expert. Any ambiguities and disagreements were listed and discussed in a session by these 2 researchers. If any problem remained after the discussion, it was investigated and resolved by the third person in the study. In the data-extracting process, the effort was to ensure that there was no missing data.

The methodological quality and potential for bias of the selected studies, particularly those evaluating diagnostic accuracy, were assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. This tool evaluates studies across 4 key domains: patient selection, index test, reference standard, and flow and timing. Each domain was assessed for risk of bias, and the first 3 domains were also evaluated for concerns regarding applicability.

## Results

This review included studies published until 2023. Through a systematic search of electronic databases and manual screening of references, 123 articles were identified. The findings of this review indicated that AI has made significant contributions to the field of medicine in various areas, including diagnosis, prognosis, and classification. Based on these findings, our included articles were classified into these three parts, presented in Tables 2 to 4.

### Application of Artificial Intelligence in the Detection of Diseases

Our review identified a total of 75 studies that employed AI algorithms for the early detection and diagnosis of medical conditions. The details of each study, including the AI model used, performance measurement, data types, and study limitations (if available), are shown in Table 2. These studies have covered a wide range of conditions, with respiratory infections, including COVID-19 disease, being the most commonly studied topic (in 18 out of the 75 studies).

The second most researched field in this category was oncology, with a focus on gastrointestinal and skin cancers.

AI was used in various applications for detection, ranging from the interpretation of medical images to the analysis of clinical and laboratory data. In this category, chest X-ray images were the most frequently used inputs (in 15 of the 75 studies), followed by CT scan images (in 9 of the 75 studies). It is worth noting that AI-based systems demonstrated high accuracy and efficiency in detecting abnormalities.

In this category, out of 75 studies reviewed, 37 studies (49.3%) have employed "deep learning methods," while 22 studies (29.3%) utilized "machine learning methods." This demonstrates a significant reliance on deep learning algorithms for the accurate detection and diagnosis of medical conditions.

### Classification

In the domain of classification, our review identified 23 studies that employed AI techniques to categorize medical data into distinct classes.

*Table 1.* Description of the PICO Components

| Population or Problem | Intervention | Comparison | Outcome |
|---|---|---|---|
| Studies of any study design that have investigated the use of AI in the diagnosis, classification, and prediction of diseases. | Any type of applied AI models | Investigating the efficacy of AI models compared to existing standard medical methods | sensitivity, specificity, accuracy, etc. |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Abbasi, 2022 (11) | machine-learning | XGBoost<br>random forest<br>SVM | COVID-19 | 6710 | chest X-ray | F1 score: 0.90<br>ROC curve: 0.96 | Pakistan | N/A |
| Abdelhamid, 2022 (12) | deep-learning | transfer-learning<br>TensorFlow `<br>optimization algorithm: RMSprop | COVID-19 | 7395 | chest X-ray | accuracy: 99.3<br>sensitivity: 99<br>specificity: 99<br>F1-Score: 99.3 | online dataset | N/A |
| Akgün, 2021 (13) | Deep Learning | VGG19, ResNet50V2, DenseNet121,<br>and MobileNet | COVID-19 | 460 | cough sound | accuracy: 86.42% | Cambridge data | N/A |
| Ali, 2022 (14) | convolutional neural networks | | COVID-19 (Omicron virus) | 915 | chest X-ray | ROC: 0.9888<br>Sensitivity: 96.2<br>Accuracy: 98<br>Precision: 100 | Egypt | N/A |
| Alotaibi, 2022 (15) | Deep Learning | Long-term and short-term memory<br><br>CNN-bidirectional LSTM<br><br>CNN-LSTM | Ischemic Stroke | 48 | MRI | F1 score: 75.3<br>Accuracy: 70.2<br>Precision: 68.1<br>F1 score: 94.2<br>Accuracy: 94.2<br>Precision: 94.5<br>F1 score: 93.8<br>Accuracy: 93.8<br>Precision: 96.5 | online dataset | N/A |
| Aponte-Hao, 2021 (16) | machine learning | XGBoost model | frailty | 5,466 | Clinical Frailty Scale (CFS) | sensitivity: 78.14<br>specificity: 74.41 | Canada | N/A |
| Babukarthik, 2022 (17) | Deep Learning | Genetic Deep Learning Convolutional Neural Network (GDCNN)<br>Particle Swarm Optimization | COVID-19 | 5071 | chest X-ray | accuracy: 97.23<br>sensitivity: 98.62<br>specificity: 97.0<br>precision: 93.0 | publicly available datasets | N/A |
| Balgetir, 2021 (18) | deep learning | VGG16, VGG19, ResNet, DenseNet,<br>MobileNet, NasNetMobile, and NasNetLarge. | multiple sclerosis | 105 | images showing plantar pressure distribution | accuracy: 89.23<br>sensitivity: 89.65<br>specificity: 88.88 | | N/A |
| Baz, 2022 (19) | deep learning | deep learning model dubbed Parallel<br>Convolution Neuron Networks (PCN2) | COVID-19 | 328 | chest X-ray | accuracy: 99.9<br>F1-score: 0.99 | online dataset | N/A |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Bendifallah,2022 (20) | Machine learning | ML models such as Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), eXtreme Gradient Boosting (XGB), and hard/soft Voting Classifier are considered ensemble learning techniques | endometriosis | 8000 | features about diagnosis, symptoms, imaging, medical treatment, fertility and surgical treatments, and follow-up. | sensitivity: 1<br>specificity: 80<br>F1-score: 88 | online dataset | N/A |
| Bhargava, 2022 (21) | machine learning | K-NN<br>SRC<br>ANN<br>SVM | COVID-19 | 31,454 | chest X-ray / CT images | Accuracy: 91.70<br>Sensitivity: 90.69<br>Specificity: 88.70<br>Accuracy: 94.40<br>Sensitivity: 72.00<br>Specificity: 86.00<br>Accuracy: 96.16<br>Sensitivity: 91.20<br>Specificity: 97.40<br>Accuracy: 99.14<br>Sensitivity: 92.86<br>Specificity: 99.86 | nine distinct datasets | N/A |
| Bozkurt, 2020 (22) | machine learning | Decision Tree, Support Vector Machines, k-Nearest Neighborhood Algorithm and Ensemble classifiers. | Obstructive Sleep Apnea | 10 | ECG | Accuracy: 85.12<br>Sensitivity: 85<br>Specificity: 86 | Turkey | N/A |
| Chen, 2019 (23) | deep learning | receptive field block, dense up sampling convolution | Prostate cancer | 50 | MRI | Recall: 90.82<br>Precision: 85.53<br>F1-score: 88.10 | dataset from MICCAI Grand Challenge | small sample size |
| Chen, 2022 (24) | multiple learning | deep attention-based MIL | Lung cancer | 1018 | Chest CT images | Recall: 87<br>accuracy: 80<br>PPV: 92<br>NPV: 59<br>AUC: 84 | Lung Image Database Consortium (LIDC-IDRI) | being depended on preexisting or human expert segmentation |
| Choi, 2021 (25) | Deep learning | Deep learning-based CAD algorithm (DCAD) | thoracic abnormalities | 244 | chest X-ray | AUC: 0.9112 | Korea | the algorithm covered only 3 major thoracic lesions |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Deperlioglu, 2020 (26) | Deep learning | Autoencoder Neural Networks (AEN) | Cardiovascular diseases | 449<br><br><br>409 | heart sounds | Accuracy: 100<br>Sensitivity: 100<br>Specificity: 100<br>Accuracy: 99.8<br>Sensitivity: 99.65<br>Specificity: 99.13 | PASCAL dataset<br><br><br>PhysioNet dataset | Potential interruptions in wireless communication and negative user experiences |
| Ding, 2021 (27) | convolutional neural network | algorithm named 'HRNet' | oral and maxillofacial disease | 912 | facial images | | China | Lower accuracy for abnormal group compared to normal group |
| Asmare, 2021 (28) | machine learning | support vector machine classifier | Rheumatic Heart Disease | 170:<br>124 cases<br>46 controls | heart sound | F1-score:96.0 ± 0.9<br>recall: 95.8 ± 1.5<br>precision: 96.2 ± 0.6<br>specificity: 96.0 ± 0.6 | in-house collected data with data from a freely available public database | N/A |
| Padmavathi Kora, 2017 (29) | Neural Network | Particle Swarm Optimization | Myocardial Infarction | 1806 | ECG Signals | 99.3% accuracy, sensitivity of 99.97%, and specificity of 98.7% | | N/A |
| Yuan Liu, 2020 (30) | deep learning | deep convolutional neural network | skin diseases | 16114 | Skin Image | accuracy: 0.66 | USA | N/A |
| Zhiyong Liu, 2021 (31) | deep learning | ResNet-101 and RPN networks | prostate cancer | | ultrasound image | | China | N/A |
| Erito Marques de Souza Filho, 2022 (32) | Machine Learning | Logistic Regression (LR), KNearest-Neighbors (KNN), Classification and Regression Tree (CART), AdaBoost (AB), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), Random Forests (RF) and Support Vector Machine (SVM) | Depression | 100 | clinical-laboratory and sociodemographic data | accuracy> .85 | Brazil | Low socio-economic information and data |
| Muhammed, M, 2021 (33) | deep learning | Alex Net is the CNN model | COVID-19 | | chest X-ray | accuracy of 97.97% | Nigeria | N/A |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Ju Gang Nam, 2019 (34) | deep learning | deep learning–based automatic detection algorithm | Malignant Pulmonary Nodules | 600 | chest X-ray | specificity of 100% | South Korean | only included malignant nodules - small nodules |
| Alvaro D. Orjuela-Cañón, 2022 (35) | Machine Learning | algorithm-in-the-loop | tuberculosis | 233 | smear sample | accuracy 78%, sensitivity 90% , specificity 33% | Colombia | high incidence of TB in the data set, selection bias |
| ChunSu Park, 2022 (36) | deep learning | N/A | bone marrow edema | 73 | MRI | sensitivity (79%) specificity (90% ) | Korea | n/A |
| Hatice Catal Reis, 2022 (37) | Deep Learning | InceptionResNetV2, InceptionV3, Mobile Net, ResNet-101, DenseNet-169, NASNetMobile, EfficientNetB0 algorithms | Covid-19 | 2400 | CT and Chest X-Ray images | 96.58 % sensitivity | Turkey | a small dataset - jpeg format |
| Prottoy Saha, 2021 (38) | Deep Neural Network | convolutional neural network(, EMCNet) | COVID-19 | 4600 | chest X-ray | 98.91% accuracy | Bangladesh | misclassify some COVID-19-positive cases as negative |
| Arkaprabha Sau, 2019 (39) | Machine Learning | CatBoost, Logistic Regression, Naïve Bayes, Random Forest | anxiety and depression | 470 | online public data | accuracy 82.6 % and precision 84.1% | India | N/A |
| Vijendra Singh, 2022 (40) | Deep Neural Network | Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic regression, Random Forest, and Naïve Bayes classifier | Chronic Kidney Disease | N/A | lab features | accuracy 100% | usa | small data sets |
| Fu M, 2018 (41) | Deep Learning | Multi-layer up-sampling structure | pancreas cancer | 236 | CT images | Precision: 77.36 ± 17.96, Recall:79.12 ± 16.27, DSC: 76.36 ± 14.34, JACCARD: 63.72 ± 17.05 | China | N/A |
| Ghazal TM, 2022 (42) | Deep Learning | AlexNet , MATLAB 2020a | Skin cancer | 2400 | image | Accuracy: 87.1% Sensitivity: 89.0% Specificity: 94.2% PPV: 93.2% NPV:82.5% | Malaysia UAE Pakistan | N/A |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Gopalakrishnan T, 2022 (43) | Deep Learning | Synergic Deep Learning (SDL) method with extreme Gradient Boosting (XGBoost), called SDL-XGBoost | Muscular dystrophy | | MRI images | Accuracy: 96.18<br>Precision:87.60<br>Sensitivity: 97.92%<br>Specificity: 95.64%<br>F-Score: 92.44<br>Kappa:0.1800 | India<br>KSA<br>Saudia Arabia<br>Yemen | N/A |
| Hemanth DJ, 2020 (44) | Deep Learning | CNN model with eight layers: image input layer, convolutional layer, ReLU layer, cross-channel normalization layer, max pooling layer, fully connected layer, softmax layer, and classifica tion layer. | diabetic retinopathy | 400 | retinal fundus images | Accuracy: 97%<br>Precision:94%<br>Sensitivity: 94%<br>Specificity: 98%<br>F-Score: 94%<br>Gmean: 96%<br>Recall: 94% | India<br>Turkey | N/A |
| Hsu FR, 2022 (45) | Deep Learning | ShuffleNet | Biliary atresia (BA) | 1976 | ultrasound images | Accuracy: 90.57%<br>Precision:85.08%<br>Sensitivity: 67.83%<br>Specificity: 96.76%<br>F-Score: 75.48%<br>AUC: 92.62% | Taiwan | the size of the US image database, the doubt of over fitting, and more requirements of test sets for verification. In addition, the overdiagnosis of BA by ShuffleNet was an issue |
| Hu S, 2020 (46) | Deep Learning | Portable Handheld Slit-Lamp Based on a Smartphone Camera | cataract | N/A | N/A | N/A | China | N/A |
| Huang C, 2022 (47) | Deep Learning | Deep Transferred EfficientNet with SVM (DTE-SVM) | Tuberculosis | N/A | CT images | Accuracy: 94.62±1.00<br>Precision: 95.30±1.24<br>Sensitivity: 93.89±1.96<br>Specificity: 95.35±1.31<br>F-Score: 94.62±1.00 | China | N/A |
| Hussain L, 2020 (48) | Machine learning | Decision Tree (DT), Naïve Bayes (NB), SVM Gaussian, SVM RBF and SVM Polynomial | congestive heart failure | 72 | RR time series interval data | Sensitivity (93.06%),<br>Specificity (81.82%),<br>Accuracy (88.79%),<br>AUC (0.95) | Saudi Arabia<br>Pakistan | N/A |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Hussain W, 2019 (49) | Machine learning | a new entropy index Permutation Fuzzy Entropy (PFEN) | Epileptic Seizure | 21 | EEG | accuracy of 98.72%, sensitivity of 98.82%, specificity of 98.63% | China Japan | N/A |
| Hwang EJ, 2019 (50) | Deep Learning | deep learning–based automatic detection algorithms (DLADs) | Active Pulmonary Tuberculosis | 65548 | chest X-ray | classification performance of 0.977–1.000, localization performance of 0.973–1.000, Sensitivities and specificities for classification were 94.3%–100% and 91.1%–100% | Korea | N/A |
| Jia X, 2022 (51) | Machine learning | WOA-SVM model | Breast Cancer | 683 | clinical data of breast cancer | Accuracy: 99% | China | N/A |
| Jia Yj, 2022 (52) | convolutional neural network | convolutional neural network VGG16 and gradient enhanced tree model | Lung Disease | | Chest sonography | Accuracy: 0.415 F-value: 0.452 Recall:0.496 | China | N/A |
| Jo Y,2017 (53) | Deep Learning | HoloConvNet | optical screening of anthrax spores | | holographic imaging | Accuracy:96.3% | Korea | N/A |
| Kaiume M, 2021 (54) | Deep Learning | a software based on a deep convolutional neural network (DCNN) | Rib fracture | 256 | CT images | Sensitivity: 0.645 (0.586–0.703) , PPV:0.793 (0.738–0.848) | Japan | N/A |
| Khan T,2019 (55) | Deep Learning | a deep learning classifier with 8 layers | Snoring | 1000 | sound | Accuracy: 0.96 | USA | N/A |
| Khurana Y, 2022 (56) | Deep Learning | ResNet-50, EfficientNetB0, VGG-16 and a custom convolutional neural network (CNN) | COVID-19 | 8000 | chest CT & X-ray | Precision:99.8% Specificity: 99.8% F1-Score: 98.7% Recall: 97.6% | India | N/A |
| N. Toda, 2023 (57) | Machine Learning | N/A | detecting pulmonary nodules, masses | 453 | chest X-ray | mean wAFROC FOM score: 93 | Japan | small and designed datasets, CT not used for ground truth labeling, US FDA guideline instead of Japanese guidelines are used. |
| S. Toften, 2021 (58) | neural network | N/A | sleep apnea | 40 | Somnofy and pulse oximetry signal | Cohen's kappa: 0.81 | Norway | small data |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| S. Toraman, 2020 (59) | convolutional neural network | N/A | COVID-19 | | chest X-ray | accuracy:97.24 | Turkey | small data |
| S. Trajanovski, 2021 (60) | deep learning | N/A | tongue squamous cell carcinoma | N/A | hyperspectral imaging (HSI) | average dice coefficient=0.891±0.053\| area under the ROC-curve=0.924±0.036 | China | N/A |
| K. E. Trinkley, 2021 (61) | Machine Learning | N/A | drug-induced long QT syndrome | 7510 | Electronic health record | N/A | United States | more analytical approaches, more types of data |
| L. M. Tseng, 2020 (62) | deep learning | N/A | Ventricular fibrillation | N/A | ECG | recall: 99<br>accuracy: 97 | Taiwan | N/A |
| A. S. Vatian, 2022 (63) | Deep Learning | N/A | acute myocardial infarction | | ECG | accuracy:85<br>F-scores : 74 | Russian | N/A |
| J. Verdu-Diaz, 2020 (64) | Machine Learning | N/A | Genetic diagnosis of muscular dystrophies | 976 | MRIs | accuracy: 95.7<br>sensitivity : 92.1<br>specificity: 99.4 | Russian | N/A |
| R. Verma, 2023 (65) | Machine Learning | N/A | glaucoma | | fundus images | Precision: 97.2<br>Recall: 97.3<br>accuracy: 97.1 | India | N/A |
| M. Viscaino, 2021 (66) | Deep Neural Network | based on convolutional and recurrent neural networks for video otoscopy analysis.Long Short-term Memory | Ear disorders | 875 | video otoscopy | accuracy: 98.15<br>precision : 91.94<br>sensitivity : 91.67<br>specificity : 98.96<br>F1-score : 91.51 | Chile | N/A |
| M. Viscaino,, 2022 (67) | convolutional neural network | N/A | otologic diagnosis | | images | accuracy: 92<br>sensitivity: 85<br>specificity: 95<br>precision: 86<br>F1-score : 85 | Chile | N/A |
| H. Wang, 2022 (68) | Deep Learning | N/A | Brown adipose tissue | 368 | PET/CT images | average DICE coefficient (DSC): 0.9057<br>average Hausdorff distance: 7.2810 | United States | small data |
| J. Wang, 2021 (69) | convolutional neural network | N/A | Congenital heart disease | 1308 | five-view echocardiograms video records | accuracy: 93.9 | china | small data |
| X. Y. Wu, 2022 (70) | Machine Learning | N/A | Coronavirus disease | 21 | CT imaging | accuracy: 59<br>Sensitivity: 91.2<br>Specificity: 18.5<br>false-positive rate: 81.5 | China | singgle center, data type, and AI model |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Y. Xiang, 2021 (71) | Deep Learning | (DenseUnet composed of encoder module , context extraction module and decoder module was established)automatic segmentation model based on the magnetic resonance image | N/A | 19 | MRI images | accuracy: 81 | China | small data |
| C. N. Xu, 2022 (72) | convolutional neural network | | glioma | 470 | MR images | Dice coefficient: 0.81<br>intraclass correlation coefficients: 0.75<br>area under the curve: 0.958 | | computational resources |
| L. Yan,, 2021 (73) | convolutional neural network | feature selection module based on attention mechanisms | cervical lesions | 1400 | colposcopy image | sensitivity: 74.6<br>accuracy: 85.5<br>specificity : 95.7<br>AUC :0.909 | China | N/A |
| Z. H. Yao, 2021 (74) | deep learning | (ResLab ) | bacterial vaginosis | | image | accuracy: 82.19 | china | N/A |
| S. L. Yi, 2022 (75) | Machine Learning | N/A | rectal cancer | 1078< | ultrasound images | accuracy: 94.66<br>precision: 94.7<br>recall: 94.65<br>F1 values : 94.67 | china | N/A |
| T. Yin, 2022 (76) | Machine Learning | upport vector machine (SVM) | functional dyspepsia | 745 | FD patients were collected from two clinical trials | accuracy: 0.773 | china | N/A |
| W. Zeng, 2022 (77) | convolutional neural network | N/A | Fetal head circumference | 999 | ultrasound images | mean absolute difference: 1.97( ± 1.89)<br>Dice similarity coefficient: 97.61 | china | N/A |
| Y. Zeng, 2020 (78) | Deep learning | N/A | colorectal cancer | 26000 | optical coherence tomography (OCT) | sensitivity: 100<br>specificity: 99.7<br>AUC: 0.998 | United States | small traing data |
| Q. Zhang, 2019 (79) | Deep learning | N/A | hyperlipidemia | 446 | physiological information and doctors' diagnosis results | accuracy: 91.49<br>sensitivity: 87.50<br>specificity: 93.33<br>precision: 87.50 | china | all factors not consider |
| X. N. Zhang, 2022 (80) | Clinical decision support systems | N/A | Type 2 diabetes | | State-University Partnership Learning Network | N/A | United States | N/A |

*Table 2.* AI in the detection and diagnosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Performance, Measurement | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Z. Zhang, 2022 (81) | Deep Learning | N/A | meibomian gland dysfunction | 2420 | Image | sensitivity: 88<br>specificity 81 | China | N/A |
| C. Zhao, 2021 (82) | convolutional neural network | N/A | Melanoma | 2420 | mei-bography images | Jaccard index: 86.84 | china | N/A |
| A. E. Zhdanov, 2023 (83) | machine learning | N/A | retinal dystrophy | | signal | N/A | Romania | N/A |
| J. Zhou, 2022 (84) | N/D | N/A | espiratory function evaluation | 220 | spirometers signal | accuracy: 68 | China | N/A |
| X. J. Zhou, 2022 (85) | convolutional neural network | N/A | dental | 210 pa-tients with one or more car-ies and 94 patients without caries | panoramic radiographs | accuracy: 0.8272<br>precision: 0.8538<br>recall: 0.8770<br>F1 score: 0.8652<br>AUC: 0.9005 | china | N/A |

*Table 3.* AI in classification of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Abdalla, 2020 (86) | conventional neural network | N/A | Arrhytmia | | ECG | | Dataset from Physionet, MIT | N/A |
| Alzubaidi, 2021 (87) | Deep Learning | N/A | breast cancer | 400 | H&E-stained breast biopsy images | Accuracy: 93.2 | BACH 2018 Grand Challenge dataset | N/A |
| | | | diabetic foot ulcers | 754 | foot skin images | precision: 95.1 recall: 94.5 F1 score: 94.8 | DFU 2 dataset | N/A |
| | | | wound types | 783 | wound images | precision: 88.1 recall: 84.8 F1 score: 86.4 accuracy: 87.94 | a combination of Google search images and DFU 2 dataset | N/A |
| Ara, 2022 (88) | Deep Learning | VGG network CNN models | Diabetic Macular Edema (DME), Choroidal Neovascularization (CNV), Drusen | 84,000 | Optical Coherence Tomography Images | Recall: 0.9990 Precision: 0.9990 F1-Score: 0.9990 Accuracy: 0.9990 | publicly available dataset | N/A |
| Badawy, 2021 (89) | fuzzy logic and deep learning | Eight CNN based SS models have been utilized: FCN-AlexNet, UNet, SegNet-VGG16, SegNet-VGG19, and DeepLabV3+(ResNet18, ResNet50, MobileNet-V2, and Xception. | breast cancer | 1200 | breast ultrasound images | accuracy: 95.45 intersection over union: 78.70 F1- score: 68.08 | MT_small_dataset. | N/A |
| Bajaj, 2017 (90) | NNLS classifier | CoHOG and Eig(Hess)-CoHOG | Alcoholism | 120 | alcoholic EEG signals | Accuracy: 95.83 Sensitivity: 100 Specificity: 91.67 | Online dataset | N/A |
| Basha, 2021 (91) | quasi-reflection-based learning procedure | swarm intelligence-driven convolutional neural network | brain tumor | 16000 | Brain MRI | accuracy: 95 | Online dataset | N/A |
| Burlina , 2017 (92) | deep learning | transfer learning and universal features derived from deep convolutional neural networks (DCNN). | age-related macular degeneration | 5664 | fundus images | accuracy: 79.4, 81.5, 93.4 for different classifications | NIH AREDS dataset | N/A |

*Table 3.* AI in classification of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Hua Li – 2019 (93) | deep learning | DenseNet-II neural network m | Breast cancer | 2042 | mammogram images | accuracy of 94.55% | China | |
| Eduardo Ramirez, 2019 (94) | neural networks and fuzzy systems | type-1 and type-2 | 2-lead cardiac arrhythmias | N/A | electrode signals or leads | 92.90% and 92.70% of classification | Mexico | N/A |
| Xinbo Ren,2023 (95) | deep learning | lock-SegNet (DBSegNet) | cardiovascular | 22210 | optical coherence tomography (OCT) images | 91.81% | China | N/A |
| Dr. K. SRIDHAR, 2022 (96) | Deep Neural Network | ResNet and DenseNet,CNN | Leukemic Cells | N/A | blood cell counts | accuracy of 95.59% | India | N/A |
| Gholami E, 2021 (97) | convolutional neural network | deep non-parametric transfer (DNPT), DeepRF | Gastric cancer | 970 | tongue images | Accuracy: 73.78 | Iran | the size of the input network |
| Gite S, 2023 (98) | deep learning | U-Net ++ | TB or other pulmonary lung diseases | 662 | X-ray images | Dice: 0.9796<br>Specificity: 0.9932<br>Mean-iou: 0.9598<br>Sensitivity: 0.9753<br>Recall: 0.9838<br>Precision:0.9685<br>Accuracy: 0.9874 | India | N/A |
| Y. Tashtoush (99) | convolutional neural networks | CNN that is augmented with convolutional block attention modules (CBAM) | Lung cancer | N/A | CT lung | acuracy: 83.49 | United States | small data |
| C. M. Vasile, 2021 (100) | deep learning ensemble method | that fused two deep learning models, one based on convolutional neural network and the other based on transfer learning (5-CNN, VGG-19) | thyroid disorders | N/A | ultrasound images | accuracy: 97.35<br>specificity: 98.43<br>Sensitivity: 95.75<br>positive predictive value: 95.41<br>negative predictive value: 98.05 | | need representative images, small data |
| W. Y. Wang, 2023 (101) | Deep Learning | PointNet++ | knee arthroplasty | N/A | Image | N/A | China | N/A |
| T. Wongsiricho, 2018 (102) | hybrid ASSC technique, Multi-Layer Hybrid Machine Learning Model | Decision Tree (DT) and Support Vector Machine | sleep disorder | 100 | polysomnographic data signal | accuracy: 0.694±0.22 in subject-specific classification and 0.942±0.02 in subject-independent classification. | Thailand | N/A |

*Table 3.* AI in classification of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| L. Yi, L. Zhang, 2022- (103) | multi-label soft-max loss (MLSL) | N/A | lung nodules | N/A | image | AUC: 82.78 | china | N/A |
| J. B. Zang, 2022 (104) | deep residual net-work model | N/A | Cardiovascular disease | 152 | ECG | accuracy : 97.89 | china | N/A |
| V. Zarikas, 2015 (105) | Bayesian net-works (BNs) | N/A | pulmonary infections | N/A | lab data | sensitivity: 90 | Austria | N/A |
| J. Zech, 2018 (106) | Natural Language Modeling | bag-of-words (BOW), word embedding, and Latent Di-richlet allocation-based ap-proaches | | 96 303 | CT reports text | AUC: 0.966 Sensitivity: 92.59 specificity : 89.67 | United States | N/A |
| X. Zhang, 2021 (107) | decision support systems with team-based care | decision support systems with team-based care | type 2 diabetes | | clinical data | online survey agree: 80 | United States | N/A |
| Q. Zhao, 2023 (108) | deep learning | N/A | | | ECG | accuracy: highly con-sistent | china | N/A |

*Table 4.* AI in determining prognosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Amyar,2022 (109) | multi-task multi-scale learning framework | multilayer perceptron and convolutional neural network (CNN) | lung cancers<br>esophageal cancers | 195 | PET images<br>PET images | area under the ROC curve:77<br>area under the ROC curve:71 | France | N/A |
| Burdick, 2020 (110) | machine learning | XGBoost Classifier | COVID-19 | 197 | clinical parameters | AUC: 0.86<br>Sensitivity: 0.90<br>Specificity: 0.58<br>LR+: 2.15<br>LR-: 0.17<br>DOR: 12.57 | US | N/A |
| chicco, 2020 (111) | Machine learning | Random forests<br>Gradient boosting<br>SVM radial | heart failure | 299 | clinical parameters | F1 score: 0.754<br>Accuracy: 0.585<br>F1 score: 0.750<br>Accuracy: 0.585<br>F1 score: 0.720<br>Accuracy: 0.543 | Pakistan | small size of the dataset |
| Bailoor , 2021 (112) | computational hemodynamic models | linear discriminant classifier | transcatheter aortic valve replacement | 29 | heart sound | accuracy: 90 | United States | N/A |
| Shaline Jia Thean Koh1- 2022 (113) | deep learning | Inception-ResNet-v2 (Gradient Class Activation Map) | Covid-19 | 795 | PCR | Accuracy:98.13%, sensitivity:97.7% specificity :99.1%. | Public open dataset | N/A |
| Renu Narain, 2016 (114) | Neural Network | quantum neural network | Cardiovascular | 689 | symptoms | 98.57% accuracy | India | N/A |
| Roberto Negro,2020 (115) | machine learning | N/A | benign Thyroid Nodule | 402 | needle aspirations | accuracy:85% - sensitivity: 0.70; specificity: 0.99 | Italian | criteria used for nodule classification - number of noduels |
| Nathan Orlando, 2020 (116) | deep learning | N/A | prostate segmentation | | ultrasound images | | | |
| Uvais Qidwai, 2022 (117) | Machine Learning-Adaptive Neuro-Fuzzy Inference System | N/A | Age-related macular-degeneration | 58 | clinical examination data | accuracy (>92%) | united kingdom | low sample size |
| Xavier Rafael-Palou, 2022 (118) | Neural Network | hierarchical generative and probabilistic network | Lung Nodule | | Lung Nodule image | accuracy of 84% | Spain | low sample size, segmentations were generated semi-automatically,our method relied on a single axial slice of the tumour |

*Table 4.* AI in determining prognosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| Göltepe Y, 2021 (119) | Machine learning | RF(Random forest); k-NN(k-nearest neighbors); NB(Naïve Bayes); LR(Logistic regression); DT(Decision tree); SVM(Support vector machine) | lung cancer | 56 | Images | Z-score Accuracy :0.83 | Turkey | N/A |
| Han X, 2021 (120) | Deep Learning | LDDMM-based Registration Network | pancreatic cancer | 40 | CT and CBCT images | improved segmentation accuracy | USA | N/A |
| Hasimbegovic E, 2021 (121) | Machine learning | ML-based approaches for understanding complex clinical decision-making processes | Severe Symptomatic Aortic Stenosis | 692 | The cohort registry data | Area under the receiver operator char acteristics curve: 0.91 Accuracy:92% Sensitivity: 92% Specificity: 90% | Austria | N/A |
| Hossain MM, 2022 (122) | Fuzzy convolutional neural network (fuzzy CNN) | QUANTITATIVE FEATURE EXTRACTION MACHINE PSO BASED FINE-TUNED FUZZY CNN | Ultrasound Image Quality Identification | 2600 | Ultrasound Image | Accuracy (99.62%), Precision (99.62%), Recall (99.61%), F1-score (99.61%) | Bangladesh Saudi Arabia Australia | N/A |
| Jung JW, 2022 (123) | Machine learning | Extreme Gradient Boosting algorithm (XGBoost) model as a machinelearning classiier | postoperative delirium following knee arthroplasty | 3980 | Clinical data | AUC score was 0.82 (95% CI: 0.80 – 0.83) and the sensitivity, speciicity was 0.72 and 0.73 respectively. | Korea | N/A |
| Işik AH, 2013 (124) | Neural Network | BP-ANN-based Mobile Information Device Applet application is developed with the Java 2 Micro Edition environment. | Chronic Pulmonary Disease | 486 | Spirometry Data | Accuracy: 98.7% , Specificity: 97.83% , Sensitivity: 97.63% , Correlation values: 0.946 | Turkey | N/A |

*Table 4.* AI in determining prognosis of medical conditions

| Study | AI Model | AI Mode details | Disease | Sample size | Data Type | AI Perfformance, Measurment | Country | limitation |
|---|---|---|---|---|---|---|---|---|
| G. Sumana, 2021 (125) | Neural Networks | N/A | renal syndromes | | laboratory data | expert evaluation assessment | India | N/A |
| J. Tang, 2023 (126) | machine learning | including Backpropagation artificial neural network (BP-ANN), random forest (RF), support vector machine (SVM), and naive Bayes classifier (NBC) | Adrenocortical carcinoma | 825 | clinical data | 5-year AUROCs=0.890, 0.847, and 0.854 | China | small data |
| A. Vodenčarevic, 2018 (127) | machine learning | N/A | Rheumatoid Arthritis (RA), a chronic inflammatory disease | | laboratory and pharmacy data | AUC value: 80 | Netherlands | N/A |
| Y. J. Wang, 2022 | Deep Learning | N/A | Alzheimer's Disease | 127 | dCDT images | accuracy: 85 | China | small data |
| G. T. Werneburg, 2022 (128) | machine learning | kernel techniques | overactive bladder treatments OnabotulinumtoxinA (OBTX-A) injection and sacral neuromodulation (SNM). | | clinical data | accuracy: 95 | United States | N/A |
| C. H. Wu, 2015 (129) | neural network models fuzzy rule-based expert system | N/A | chronic kidney disease | | clinical and lab data | accuracy: 88.40 | | N/A |
| I. Y. Zhang, 2023 (130) | machine learning | N/A | malignant tumor | | clinical value | (for pediatric, adolescent, and young adult) average C-index = 86.8%, 85.2%, and 88.6% \|average time-dependent AUC = 76.5%, 88.1%, and 99.0% | United States | not available data, |
| C. J. Zimmermann, 2021 (131) | feedback and adapted the tool | N/A | ill older adult trauma | 48 | trauma clinicians in Wisconsin, Texas, and Oregon | qualitative content analysis | United States | N/A |
| A. Vallée, 2022 (132) | N/A | N/A | human immunodeficiency virus (HIV) | 8180 | electronic medical record (EMR) | N/A | France | A patient with a previous HIV diagnosis was included, did not compare rates of HIV testing during the study |

Table 3 presents the details of each study. These studies covered a wide range of medical conditions, with oncology (specifically breast cancer) being the most studied topic in 7 of 23 studies. Another widely researched area was cardiovascular diseases, which were addressed in 4 of 23 studies. AI algorithms demonstrated remarkable capabilities in accurately classifying patients based on various data, with breast images (ultrasound or mammograms) being the most used items.

Out of 23 articles reviewed in this category, 11 studies (47.8%) have utilized “deep learning methods.” In addition, 4 studies (17.4%) have utilized “neural networks,” further highlighting the versatility and effectiveness of AI in classifying medical conditions.

### Prognosis and Prevention

In the domain of prognosis and prevention, our review identified 25 studies that leveraged AI methodologies to predict the progression and outcomes of medical conditions, as shown in Table 4. These studies covered a wide spectrum of diseases, ranging from heart failure to neurological disorders. Oncology again was the most studied field in this category. AI-based prognostic models exhibited impressive predictive performance, enabling clinicians to anticipate patient outcomes with greater accuracy and foresight. The category's most frequently used items were clinical and laboratory parameters, including PCR, used in 12 out of 25 articles. Furthermore, AI algorithms demonstrated the ability to identify prognostic factors that might otherwise go unnoticed, thereby facilitating early intervention and risk mitigation strategies.

Out of 25 articles reviewed in this category, 14 studies (56%) have utilized “machine learning methods” for prognosis and prevention tasks. However, deep learning methods were only employed in 4 articles (16%).

### Comparing the 3 Classes

Furthermore, we evaluated research in the 3 classes based on sample size, AI models, and data type.

Figure 2 shows the comparison from the aspect of sample size. As the result shows for the classification, the AI-based system required a larger sample size. Also, detection and diagnosis need a larger sample size than prognosis. Considering the AI models, we first investigated the AI-based model as presented in Figure 3. The most used model is DL. Then, we analyzed them in each of the 3 classes. The most used model in detection-diagnosis and classification is the deep learning model, with about 49.3% and 47.8 %, respectively, and prognosis uses ML, with 56%. These results indicate that for detection and classification, deep learning is the most used model, while in prognosis, in which there is a need to find patterns among data, ML is the most used model that uses a smaller sample size than the deep learning model.

We further evaluated the classes from the aspect of data type in AI-based systems. As the results in Tables 2 to 4 indicate, in detection-diagnosis and classification, different modalities of images and signals are the 2 most used data types. More than 75% and 60% are images, and more than 15% and 26 % are signals in those 2 classes. Also, the most used data type in prognosis is laboratory data, with more than 53% usage. These results show that for detection and classification, images and signals carry proper information,

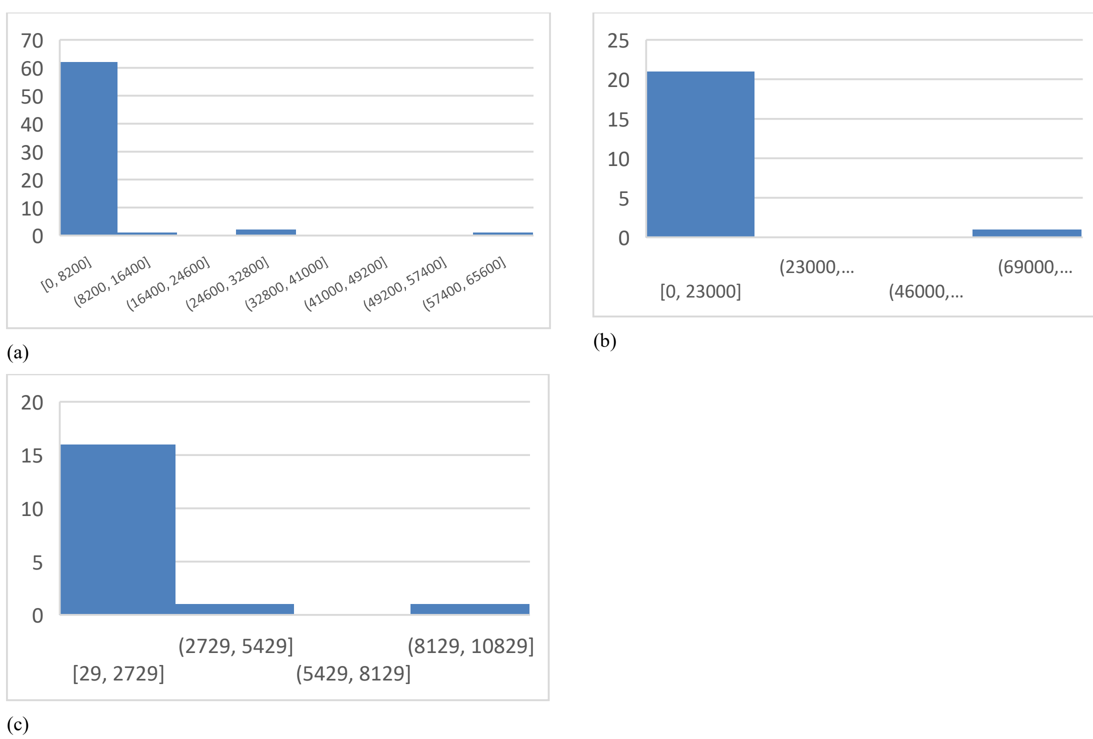


*Figure 2.* (A) Detection and Diagnosis, (B) classification, and (C) prognosis

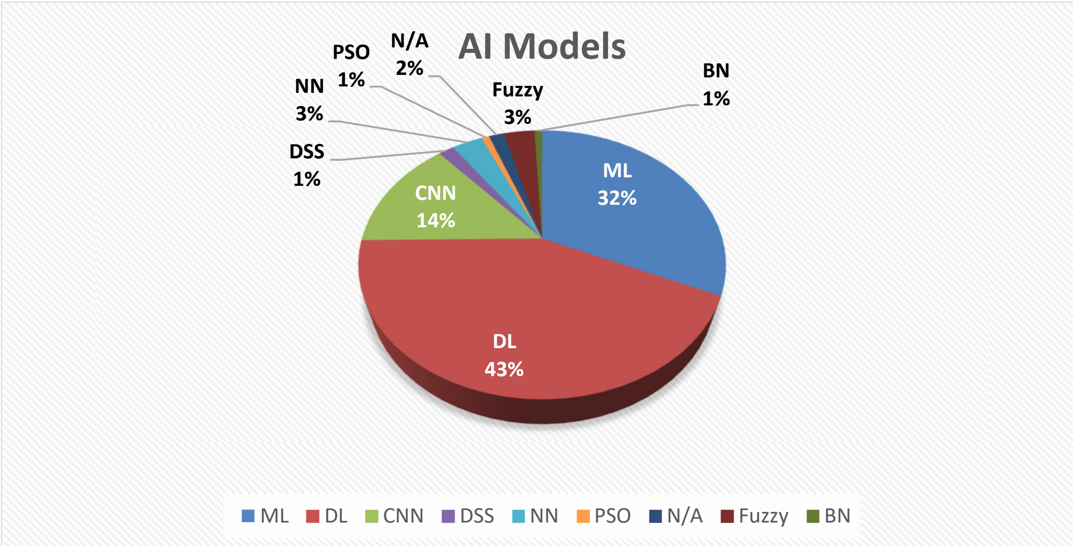


*Figure 3.* AI-based models in studies. Neural networks (NN), convolutional neural networks (CNN), and deep learning (DL) stand for 3 types of network-based ML algorithms, and ML class stands for the rest of the ML algorithms.

*Table 5.* Methodology checklist: the QUADAS-2 tool components for risk of bias assessment the selected studies (questions)

| |
|---|
| 1.Was a consecutive or random sample of patients enrolled? (Yes / No / Unclear) |
| 2.Was a case-control design avoided? (Yes / No / Unclear) |
| 3.Did the study avoid inappropriate exclusions? (Yes / No / Unclear) |
| 4.Could the selection of patients have introduced bias? (Risk: Low / High / Unclear) |
| 5.Is there concern that the included patients do not match the review question? (Concern: Low / High / Unclear) |
| 6.Were the index test results interpreted without knowledge of the results of the reference standard? (Yes / No / Unclear) |
| 7.If a threshold was used, was it pre-specified? (Yes / No / Unclear) |
| 8.Could the conduct or interpretation of the index test have introduced bias? (Risk: Low / High / Unclear) |
| 9.Is there concern that the index test, its conduct, or interpretation differ from the review question? (Concern: Low / High / Unclear) |
| 10.Is the reference standard likely to correctly classify the target condition? (Yes / No / Unclear) |
| 11.Were the reference standard results interpreted without knowledge of the results of the index test? (Yes / No / Unclear) |
| 12.Could the reference standard, its conduct, or its interpretation have introduced bias? (Risk: Low / High / Unclear) |
| 13.Is there concern that the target condition as defined by the reference standard does not match the review question? (Concern: Low / High / Unclear) |
| 14.Was there an appropriate interval between index test(s) and reference standard? (Yes / No / Unclear) |
| 15.Did all patients receive a reference standard? (Yes / No / Unclear) |
| 16.Did patients receive the same reference standard? (Yes / No / Unclear) |
| 17.Were all patients included in the analysis? (Yes / No / Unclear) |
| 18.Could the patient flow have introduced bias? (Risk: Low / High / Unclear) |

while in prognosis, laboratory data analysis leads to extracting early signs of disease.

The specific signaling questions used for the methodological quality and potential for bias of the selected studiesassessment, adapted from the standard QUADAS-2 checklist, are provided in Table 5. The summary of the risk of bias and applicability judgments for the diagnostic accuracy studies included in this review (as listed in Table 2) is presented in Figure 4.

## Discussion

The information obtained from studies was analyzed from several aspects, including the applied AI models, the investigated disease, sample size, data type, and measurement criteria. Most studies used structures based on DL in the model section. The most used methods were deep learning, ML, CNN, hybrid fuzzy-based learning systems, classical neural networks (NN), decision support systems (DSS), Bayesian network (BN), and particle swarm optimization (PSO). ML algorithms consist of two classes of non-network-based ML algorithms, such as random forest, and network-based ML algorithms, such as NN. Here, for precise classification, we separate each type of network-based algorithm of NN, CNN, and DL, which are mostly used in papers, and consider the rest of the ML algorithms as ML class.

From the point of view of diseases, the two top investigated diseases were COVID-19 and cancer. Also, in terms of data type, various medical data types such as images, signals, clinical data, and geographical data were used in studies. However, medical images were used as the data investigated in most of the studies.

In the sample size section, the largest data volume included 96,303 CT image samples, and the smallest included 19 MRI image samples. It shows that the proper sample size

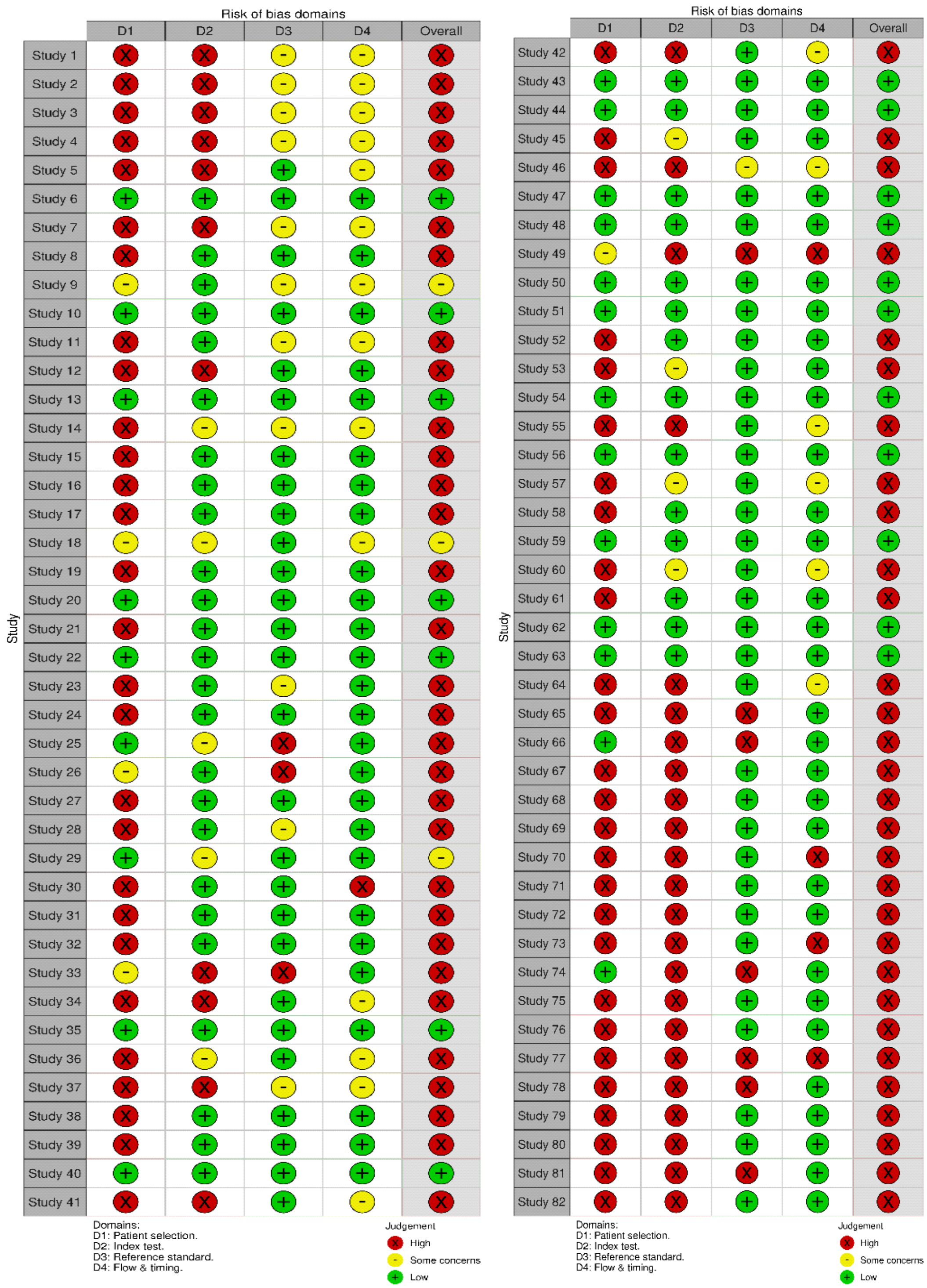


*Figure 4.* Methodology checklist: the QUADAS-2 tool for studies of diagnostic test accuracy (listed according to table 2)

depends on the applied model. Besides, it should be noted that this sample size variation might affect results. Moreover, on the measurement criteria, accuracy, specificity, F1-score, and area under the curve were among the most used criteria in studies, and the accuracy of the model was at the top. The highest accuracy value of the models is equal to 99.9% in the diagnosis of COVID-19 using deep learning and 328 data points related to chest radiography, and also in the classification of diabetic macular edema using deep learning and 84000 tomographic images.

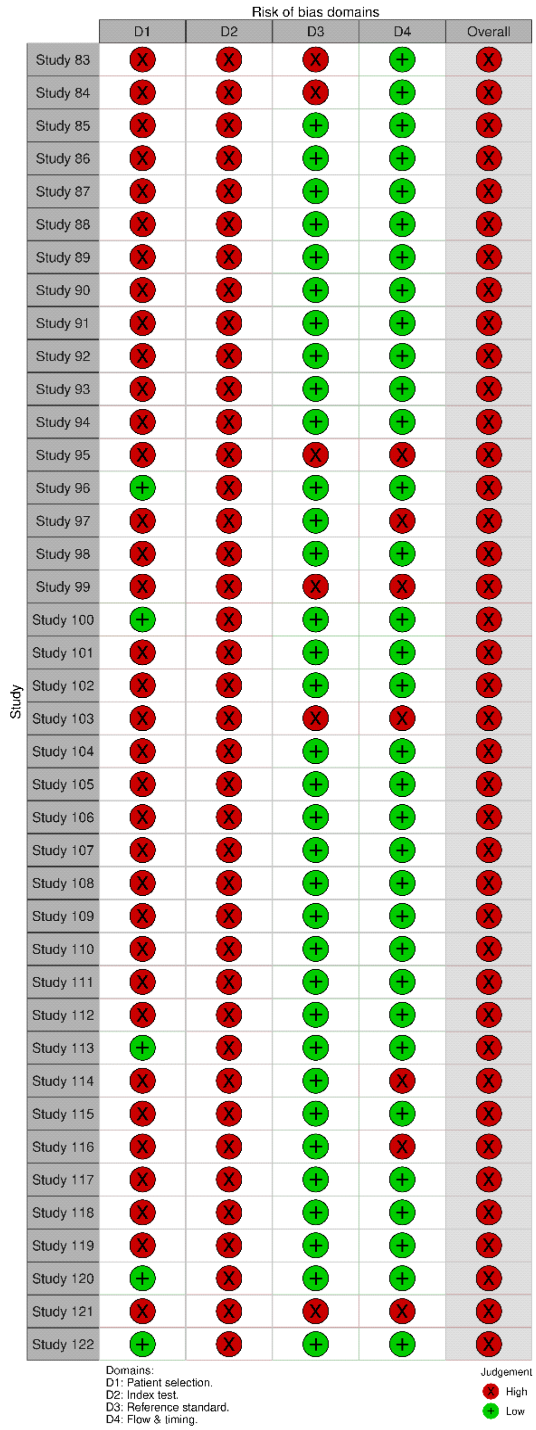


*Figure 4.* Methodology checklist: the QUADAS-2 tool for studies of diagnostic test accuracy (listed according to table 2)

### Architectural Suitability and Domain Characteristics

CNNs are particularly well-suited to high-dimensional imaging tasks because their convolutional filters and pooling layers efficiently learn and aggregate spatial features, such as edges, textures, and anatomical structures, across multiple scales. In contrast, RNNs (and LSTM variants) excel at modeling temporal dependencies in sequential data, such as time-series laboratory values or vital signs, by maintaining an internal state that captures information

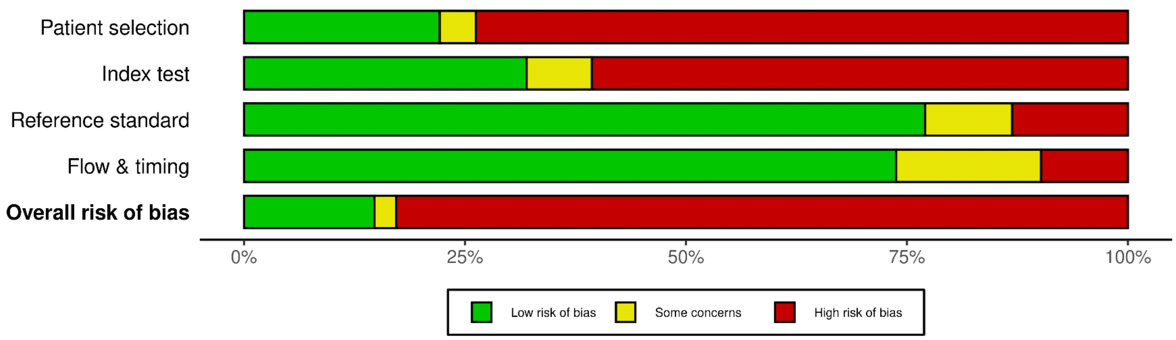


*Figure 4.* Methodology checklist: the QUADAS-2 tool for studies of diagnostic test accuracy (listed according to table 2)

across prior time steps. The nature of the input modality further dictates model choice and hyperparameter tuning: structured clinical data (eg, tabular lab results, demographics) often employ fully connected networks or tree-based models with feature engineering and regularization parameters (eg, learning rate, tree depth) optimized for tabular distributions, whereas unstructured imaging data require convolutional architectures with appropriately chosen filter sizes, depths, and spatial dropout rates to balance representational power and overfitting risk. Finally, hybrid CNN-RNN pipelines enable multimodal integration—first extracting spatial embeddings from images via CNNs, then modeling their temporal evolution or combining them with sequential clinical measurements in RNN layers—thereby capturing both spatial and temporal patterns for more comprehensive clinical predictions.

### Limitations

A key strength of this study is its comprehensive and systematic approach, as it analyzed a wide range of studies from multiple reputable databases, ensuring a thorough review of AI applications in medical diagnosis, classification, and prognosis. Additionally, the study highlighted AI's significant contributions across different medical fields, particularly in disease detection and prediction, providing valuable insights for future research. However, a notable weakness is the lack of emphasis on the speed and efficiency of AI models, as the study primarily focuses on accuracy without discussing the practical implications of AI adoption in clinical settings. Another limitation is its exclusion of non-English studies and studies with inaccessible full texts, which may introduce selection bias and limit the generalizability of its findings. Furthermore, while the paper categorizes AI applications effectively, it does not critically assess the challenges of AI implementation, such as ethical concerns, data biases, and real-world integration hurdles. Finally, considering the interpretability and data requirements, we should explain that DL models, compared to ML models, are less interpretable and need more data. However, using explainable artificial intelligence (XAI) techniques, this problem can be solved. Also, the mentioned limitations in the studies include lack of data, having small datasets that can be solved using data augmentation or federated learning, asymmetry of data, dependence on the data labeling process, incomplete or inaccurate data that lead to data bias, and the need for an expert's opinion in the data labeling process.

Moreover, many high-performance DL models remain “black boxes,” limiting clinician trust; integrating XAI methods such as saliency mapping or SHAP can partially mitigate this but adds complexity. The reliance on English-language, publicly accessible datasets may also introduce geographic and demographic biases, and small sample sizes exacerbate overfitting; strategies like data augmentation and federated learning could improve generalizability. Practical integration into clinical workflows remains challenging due to workflow disruption, interoperability issues with electronic health records, and regulatory uncertainties, underscoring the need for clinician-AI–AI codesign. Finally, ethical and regulatory concerns, including patient privacy, accountability for AI-driven decisions, and the lack of standardized approval pathways—must be addressed to ensure safe and compliant deployment.

### Conclusion and Future Work

AI-based methods with a variety of approaches can be used in different areas of medicine, such as automated triage systems or real-time imaging analysis, based on the type of data and the amount of available data. DL, as the most popular strategy, has multiple processing layers in the network. They are used in many medical applications with high accuracy. In medical diagnoses, classification, and prognoses, the focus has been on the use of AI methods, which increase the accuracy, and criteria for selecting methods that increase the speed of diagnoses, classification, and prognoses were not provided. This may be a suggestion for future investigation. Also, the efficiency of AI-based methods compared with manual ones is discussed in most studies, which indicates the willingness to use AI in the medical industry. Also, with the human in the loop in medicine and medical staff duty, all AI-based applications are assistants, and the final choice is made by the human expert who uses such AI-based aid.

As a result, no work is focused on replacing clinical professionals with AI. This demonstrates the complementarity and aid of artificial intelligence in medical services. It is indicated that in the future, it will be explored in which medical services AI may be fully implemented, as well as how human presence will be involved. Robotic surgery is one relevant example in this subject. Robotic surgery is one relevant example in this subject.

### Authors' Contributions

M.N. and E.B. conceptualized the study. N.R.R., Z.R., and M.R.B. conducted the literature search and data extraction. N.R.R. and M.S. contributed to data analysis and manuscript revision. All authors reviewed and approved the final manuscript.

### Ethical Considerations

This study was approved by the Ethics Committee of Iran University of Medical Sciences (IR.IUMS.REC.1401.711).

### Acknowledgment

The authors would like to thank the Iran University of Medical Sciences for their support.

### Conflict of Interests

The authors declare that they have no competing interests.

## References

1. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, et al. Artificial intelligence: A powerful paradigm for scientific research. Innovation. 2021;2(4):100179.
2. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6(2):94-8.
3. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23(1):689.
4. Elter M, Schulz-Wendtland R, Wittenberg T. The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process. Med Phys. 2007;34(11):4164-72.
5. Gholami E TS, Kheirabadi M. Diagnosis of Gastric Cancer via Classification of the Tongue Images using Deep Convolutional Networks. J Inf Syst Telecommun (JIST). 2021;35(9):191-6.
6. Gu ZC, Huang SR, Dong L, Zhou Q, Wang J, Fu B, et al. An Adapted Neural-Fuzzy Inference System Model Using Preprocessed Balance Data to Improve the Predictive Accuracy of Warfarin Maintenance Dosing in Patients After Heart Valve Replacement. Cardiovasc Drugs Ther. 2022;36(5):879-89.
7. Gite S, Mishra A, Kotecha K. Enhanced lung image segmentation using deep learning. Neural Comput Appl. 2023;35(31):22839-53.
8. Jia Y. Application of Recurrent Neural Network Algorithm in Intelligent Detection of Clinical Ultrasound Images of Human Lungs. Comput Intell Neurosci. 2022;2022:9602740.
9. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J. 2021;8(2):e188-e94.
10. PRISMA Flow Diagram Generator 2019 [Available from: http://prisma.thetacollaborative.ca/.
11. Abbasi WA, Abbas SA, Andleeb S, Bibi M, Majeed F, Jaleel A, et al. COVIDX: Computer-aided diagnosis of COVID-19 and its severity prediction with raw digital chest X-ray scans. Quant Biol. 2022;10(2):208-20.
12. Abdelhamid AA, Abdelhalim E, Mohamed MA, Khalifa F. Multi-Classification of Chest X-rays for COVID-19 Diagnosis Using Deep Learning Algorithms. Appl Sci (Switzerland). 2022;12(4).
13. Akgün D, Kabakuş AT, Karapinar Şentürk Z, Şentürk A, Küçükkülahli E. A transfer learning-based deep learning approach for automated COVID-19 diagnosis with audio data. Turk J Electr Eng Comput Sci. 2021;29:2807-23.
14. Ali EM, Elnady MA, Abdella AK, Elstohy RA. Detection of omicron virus based on convolutional neural network. J Theor Appl Inf Technol. 2022;100(9):2871-84.
15. Alotaibi NS, Alotaibi AS, Eliazer M, Srinivasulu A. Detection of Ischemic Stroke Tissue Fate from the MRI Images Using a Deep Learning Approach. Mob Inf Syst. 2022;2022.
16. Aponte-Hao S, Wong ST, Thandi M, Ronksley P, McBrien K, Lee J, et al. Machine learning for identification of frailty in Canadian primary care practices. Int J Popul Data Sci. 2021;6(1).
17. Babukarthik RG, Chandramohan D, Tripathi D, Kumar M, Sambasivam G. COVID-19 identification in chest X-ray images using intelligent multi-level classification scenario. Comput Electr Eng. 2022;104.
18. Balgetir F, Bilek F, Kakakus S, Arslan-Tuncer S, Demir CF. Detection of ataxia in low disability MS patients by hybrid convolutional neural networks based on images of plantar pressure distribution. Mult Scler Relat Disord. 2021;56.
19. Baz A, Baz M. Pcn2: Parallel cnn to diagnose covid-19 from radiographs and metadata. Intell Autom Soft Comput. 2022;31(2):1051-69.
20. Bendifallah S, Puchar A, Suisse S, Delbos L, Poilblanc M, Descamps P, et al. Machine learning algorithms as new screening approach for patients with endometriosis. Sci Rep. 2022;12(1).
21. Bhargava A, Bansal A, Goyal V. Machine learning-based automatic detection of novel coronavirus (COVID-19) disease. Multimed Tools Appl. 2022;81(10):13731-50.
22. Bozkurt F, Ucar MK, Bozkurt MR, Bilgin C. Detection of Abnormal Respiratory Events with Single Channel ECG and Hybrid Machine Learning Model in Patients with Obstructive Sleep Apnea. IRBM. 2020;41(5):241-51.
23. Chen A, Zhu L, Zang H, Ding Z, Zhan S. Computer-aided diagnosis and decision-making system for medical data analysis: A case study on prostate MR images. J Manag Sci Eng. 2019;4(4):266-78.
24. Chen JH, Zeng HY, Zhang C, Shi ZW, Dekker A, Wee L, et al. Lung cancer diagnosis using deep attention-based multiple instance learning and radiomics. Med Phys. 2022;49(5):3134-43.
25. Choi SY, Park S, Kim M, Park J, Choi YR, Jin KN. Evaluation of a deep learning-based computer-aided detection algorithm on chest radiographs: Case-control study. Medicine. 2021;100(16):e25663.
26. Deperlioglu O, Kose U, Gupta D, Khanna A, Sangaiah AK. Diagnosis of heart diseases by a secure Internet of Health Things system based on Autoencoder Deep Neural Network. Comput Commun. 2020;162:31-50.
27. Ding M, Kang Y, Yuan Z, Shan X, Cai Z. Detection of facial landmarks by a convolutional neural network in patients with oral and maxillofacial disease. Int J Oral Maxillofac Surg. 2021;50(11):1443-9.
28. Asmare MH, Filtjens B, Woldehanna F, Janssens L, Vanrumste B. Rheumatic heart disease screening based on phonocardiogram. Sensors. 2021;21(19).
29. Kora P. ECG based Myocardial Infarction detection using Hybrid Firefly Algorithm. Comput Methods Programs Biomed. 2017;152:141-8.
30. Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, et al. A deep learning system for differential diagnosis of skin diseases. Nat Med. 2020;26(6):900-8.
31. Liu Z, Yang C, Huang J, Liu S, Zhuo Y, Lu X. Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer. Future Gener Comput Syst. 2021;114:358-67.
32. de Souza Filho EM, Veiga Rey HC, Frajtag RM, Arrowsmith Cook DM, Dalbonio de Carvalho LN, Pinho Ribeiro AL, et al. Corrigendum to "Can machine learning be useful as a screening tool for depression in primary care?" J Psychiatr Res. 2022 Apr:148:61-62.
33. Muhammed M, Boukar MM, Aldullahi SE, Dane S. The Application of Artificial Intelligence Technique (CNN-Alexnet) in Diagnosing COVID-19 Using Chest X-ray Images. J Res Med Dent Sci. 2021;9(5):21-6.
34. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. Radiology. 2019;290(1):218-28.
35. Orjuela-Canon AD, Jutinico AL, Awad C, Vergara E, Palencia A. Machine learning in the loop for tuberculosis diagnosis support. Front Public Health. 2022;10.
36. Park C, Kim M, Park C, Son W, Lee SM, Seok Jeong H, et al. Diagnostic performance for detecting bone marrow edema of the hip on

dual-energy CT: Deep learning model vs. musculoskeletal physicians and radiologists. Eur J Radiol. 2022;152.
37. Reis HC, Turk V. COVID-DSNet: A novel deep convolutional neural network for detection of coronavirus (SARS-CoV-2) cases from CT and Chest X-Ray images. Artif Intell Med. 2022;134.
38. Saha P, Sadi MS, Islam MM. EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. Informatics in Medicine Unlocked. 2021;22.
39. Sau A, Bhakta I. Erratum: Screening of anxiety and depression among seafarers using machine learning technology (Informatics in Medicine Unlocked (2019) 16, (S235291481830193X), (10.1016/j.imu.2018.12.004)). Inform Med Unlocked. 2019;16.
40. Singh V, Asari VK, Rajasekaran R. A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. Diagnostics. 2022;12(1).
41. Fu M, Wu W, Hong X, Liu Q, Jiang J, Ou Y, et al. Hierarchical combinatorial deep learning architecture for pancreas segmentation of medical computed tomography cancer images. BMC Syst Biol. 2018;12.
42. Ghazal TM, Hussain S, Khan MF, Khan MA, Said RAT, Ahmad M. Detection of Benign and Malignant Tumors in Skin Empowered with Transfer Learning. Computational Intelligence and Neuroscience. 2022;2022.
43. Gopalakrishnan T, Sudhakaran P, Ramya KC, Sathesh Kumar K, Al-Wesabi FN, Alohali MA, et al. An automated deep learning based muscular dystrophy detection and classificationmodel. Comput Mater Continua. 2022;71(1):305-20.
44. Hemanth DJ, Deperlioglu O, Kose U. An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. Neural Computing and Applications. 2020;32(3):707-21.
45. Hsu FR, Dai ST, Chou CM, Huang SY. The application of artificial intelligence to support biliary atresia screening by ultrasound images: A study based on deep learning models. PLoS One. 2022;17(10 October).
46. Hu S, Wu H, Luan X, Wang Z, Adu M, Wang X, et al. Portable Handheld Slit-Lamp Based on a Smartphone Camera for Cataract Screening. J Ophthalmol. 2020;2020.
47. Huang C, Wang W, Zhang X, Wang SH, Zhang YD. Tuberculosis Diagnosis using Deep Transferred EfficientNet. IEEE/ACM Trans Comput Biol Bioinform. 2022;Pp.
48. Hussain L, Aziz W, Khan IR, Alkinani MH, Alowibdi JS. Machine learning based congestive heart failure detection using feature importance ranking of multimodal features. Math Biosci Eng. 2020;18(1):69-91.
49. Hussain W, Wang B, Niu Y, Gao Y, Wang X, Sun J, et al. Epileptic Seizure Detection With Permutation Fuzzy Entropy Using Robust Machine Learning Techniques. IEEE ACCESS. 2019;7:182238-58.
50. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. Clin Infect Dis. 2019;69(5):739-47.
51. Jia X, Sun XL, Zhang XA. Breast Cancer Identification Using Machine Learning. Math Probl Eng. 2022;2022.
52. Jia YJ. Application of Recurrent Neural Network Algorithm in Intelligent Detection of Clinical Ultrasound Images of Human Lungs. Comput Intellig Neurosci. 2022;2022.
53. Jo Y, Park S, Jung J, Yoon J, Joo H, Kim MH, et al. Holographic deep learning for rapid optical screening of anthrax spores. Sci Adv. 2017;3(8).
54. Kaiume M, Suzuki S, Yasaka K, Sugawara H, Shen Y, Katada Y, et al. Rib fracture detection in computed tomography images using deep convolutional neural networks. Medicine (Baltimore). 2021;100(20):E26024.
55. Khan T. A deep learning model for snoring detection and vibration notification using a smartwearable gadget. Electronics (Switzerland). 2019;8(9).
56. Khurana Y, Soni U. Leveraging deep learning for COVID-19 diagnosis through chest imaging. Neural Computing and Applications. 2022;34(16):14003-12.
57. Toda N, Hashimoto M, Iwabuchi Y, Nagasaka M, Takeshita R, Yamada M, et al. Validation of deep learning-based computer-aided detection software use for interpretation of pulmonary abnormalities on chest radiographs and examination of factors that influence readers' performance and final diagnosis. Jpn J Radiol. 2023;41(1):38-44.
58. Toften S, Kjellstadli JT, Tyvold SS, Moxness MHS. A Pilot Study of Detecting Individual Sleep Apnea Events Using Noncontact Radar Technology, Pulse Oximetry, and Machine Learning. J Sens. 2021;2021.
59. Toraman S, Alakus TB, Turkoglu I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. Chaos Solitons Fractals. 2020;140.
60. Trajanovski S, Shan C, Weijtmans PJC, De Koning SGB, Ruers TJM. Tongue Tumor Detection in Hyperspectral Images Using Deep Learning Semantic Segmentation. IEEE Trans Biomed Eng. 2021;68(4):1330-40.
61. Trinkley KE, Pell JM, Martinez DD, Maude NR, Hale G, Rosenberg MA. Assessing Prescriber Behavior with a Clinical Decision Support Tool to Prevent Drug-Induced Long QT Syndrome. Appl Clin Inform. 2021;12(1):190-7.
62. Tseng LM, Tseng VS. Predicting Ventricular Fibrillation through Deep Learning. IEEE Access. 2020;8:221886-96.
63. Vatian AS, Gusarova NF, Dobrenko NV, Zmievsky DA, Kabyshev MV, Polevaya TA, et al. Automated evaluation of ECG parameters during the COVID-19 pandemic. Sci Tech Inf Technol.Mech Opt. 2022;22(6):1166-77.
64. Verdu-Diaz J, Alonso-Perez J, Nunez-Peralta C, Tasca G, Vissing J, Straub V, et al. Accuracy of a machine learning muscle MRI-based tool for the diagnosis of muscular dystrophies. Neurology. 2020;94(10):E1094-E102.
65. Verma R, Shrinivasan L, Hiremath B. Machine learning classifiers for detection of glaucoma. Int J Artif Intell (IAES). 2023;12(2):806-14.
66. Viscaino M, Maass JC, Delano PH, Cheein FA. Computer-Aided Ear Diagnosis System Based on CNN-LSTM Hybrid Learning Framework for Video Otoscopy Examination. IEEE Access. 2021;9:161292-304.
67. Viscaino M, Talamilla M, Maass JC, Henríquez P, Délano PH, Cheein CA, et al. Color Dependence Analysis in a CNN-Based Computer-Aided Diagnosis System for Middle and External Ear Diseases. Diagnostics. 2022;12(4).
68. Wang H, Wang Z, Wang J, Li K, Geng G, Kang F, et al. ICA-Unet: An improved U-net network for brown adipose tissue segmentation. J Innov Opt Health Sci. 2022;15(3).
69. Wang J, Liu X, Wang F, Zheng L, Gao F, Zhang H, et al. Automated interpretation of congenital heart disease from multi-view echocardiograms. Med Image Anal. 2021;69.
70. Wu XY, Ding F, Li K, Huang WC, Zhang Y, Zhu J. Analysis of the Causes of Solitary Pulmonary Nodule Misdiagnosed as Lung Cancer by Using Artificial Intelligence: A Retrospective Study at a Single Center. Diagnostics. 2022;12(9).
71. Xiang Y, Wu Y, Zhang X, Hu X, Liu J, Lei L, et al. Automatic segmentation of levator ani muscle in MRI images based on DenseUnet model. Journal of Third Military Medical University. 2021;43(18):1720-8.
72. Xu CN, Peng YY, Zhu WF, Chen ZY, Li JR, Tan WH, et al. An automated approach for predicting glioma grade and survival of LGG patients using CNN and radiomics. Front Oncol. 2022;12.
73. Yan L, Li S, Guo Y, Ren P, Song H, Yang J, et al. Multi-state colposcopy image fusion for cervical precancerous lesion diagnosis using BF-CNN. Biomed Signal Process Control. 2021;68.
74. Yao ZH, Chen W, Li C, Yang HY, He YL, Tan YS, et al. Automatic Diagnosis of Vaginal Microecological Pathological Images Based on Deep Learning. Prog Biochem Biophys. 2021;48(11):1348-57.
75. Yi SL, Wei YR, Luo XM, Chen D. Diagnosis of rectal cancer based on the Xception-MS network. Phys Med Biol. 2022;67(19).
76. Yin T, Zheng H, Ma T, Tian X, Xu J, Li Y, et al. Predicting acupuncture efficacy for functional dyspepsia based on routine clinical features: a machine learning study in the framework of predictive, preventive, and personalized medicine. EPMA J. 2022;13(1):137-47.
77. Zeng W, Luo J, Cheng JR, Lu YL. Efficient fetal ultrasound image segmentation for automatic head circumference measurement using a lightweight deep convolutional neural network. MEDICAL PHYSICS. 2022;49(8):5081-92.
78. Zeng Y, Xu S, Chapman WC, Li S, Alipour Z, Abdelal H, et al. Real-time colorectal cancer diagnosis using PR-OCT with deep learning. Theranostics. 2020;10(6):2587-96.
79. Zhang Q, Liu YL, Liu GH, Zhao G, Qu ZG, Yang WM. An automatic diagnostic system based on deep learning, to diagnose hyperlipidemia. Diabetes Metab Syndr. 2019;12:637-45.
80. Zhang XN, Svec M, Tracy R, Ozanich G. Clinical decision support systems with team-based care on type 2 diabetes improvement for

Medicaid patients: A quality improvement project. Int J Med Inform. 2022;158.
81. Zhang Z, Lin X, Yu X, Fu Y, Chen X, Yang W, et al. Meibomian Gland Density: An Effective Evaluation Index of Meibomian Gland Dysfunction Based on Deep Learning and Transfer Learning J Clin Med. 2022;11(9).
82. Zhao C, Shuai R, Ma L, Liu WJ, Wu ML. Segmentation of dermoscopy images based on deformable 3D convolution and ResU-NeXt plus. Med Biol Eng Comput. 2021;59(9):1815-32.
83. Zhdanov AE, Dolganov AY, Zanca D, Borisov VI, Luchian E, Dorosinsky LG. Evaluation of the effectiveness of the decision support algorithm for physicians in retinal dystrophy using machine learning methods. Comput Opt. 2023;47(2):272-7.
84. Zhou J, Wang P, Guo L, Cao J, Zhou M, Dai R. Automated interpretation of the pulmonary function test by a portable spirometer in Chinese adults. Clin Respir J. 2022;16(8):555-61.
85. Zhou XJ, Yu GX, Yin QY, Liu Y, Zhang ZL, Sun J. Context Aware Convolutional Neural Network for Children Caries Diagnosis on Dental Panoramic Radiographs. Comput Math Methods Med. 2022;2022.
86. Abdalla FYO, Wu L, Ullah H, Ren G, Noor A, Mkindu H, et al. Deep convolutional neural network application to classify the ECG arrhythmia. Signal Image Video Process. 2020;14(7):1431-9.
87. Alzubaidi L, Fadhel MA, Al-Shamma O, Zhang JL, Santamaria J, Duan Y. Robust application of new deep learning tools: an experimental study in medical imaging. Multimed Tools Appl. 2022;81(10):13289-317.
88. Ara RK, Matiolański A, Dziech A, Baran R, Domin P, Wieczorkiewicz A. Fast and Efficient Method for Optical Coherence Tomography Images Classification Using Deep Learning Approach. Sensors. 2022;22(13).
89. Badawy SM, Mohamed AEA, Hefnawy AA, Zidan HE, GadAllah MT, El-Banby GM. Automatic semantic segmentation of breast tumors in ultrasound images based on combining fuzzy logic and deep learning-A feasibility study. PloS One. 2021;16(5).
90. Bajaj V, Guo YH, Sengur A, Siuly S, Alcin OF. A hybrid method based on time-frequency images for classification of alcohol and control EEG signals. Neural Comput. Appl. 2017;28(12):3717-23.
91. Basha J, Bacanin N, Vukobrat N, Zivkovic M, Venkatachalam K, Hubalovsky S, et al. Chaotic Harris Hawks Optimization with Quasi-Reflection-Based Learning: An Application to Enhance CNN Design. SENSORS. 2021;21(19).
92. Burlina P, Pacheco KD, Joshi N, Freund DE, Bressler NM. Comparing humans and deep learning performance for grading AMD: A study in using universal deep features and transfer learning for automated AMD analysis. Comput Biol Med. 2017;82:80-6.
93. Li H, Zhuang S, Li DA, Zhao J, Ma Y. Benign and malignant classification of mammogram images based on deep learning. Biomedical Signal Processing and Control. 2019;51:347-54.
94. Ramirez E, Melin P, Prado-Arechiga G. Hybrid model based on neural networks, type-1 and type-2 fuzzy systems for 2-lead cardiac arrhythmia classification. Expert Syst Appl. 2019;126:295-307.
95. Ren X, Zhao Y, Fan J, Wu H, Chen Q, Kubo T. Semantic segmentation of superficial layer in intracoronary optical coherence tomography based on cropping-merging and deep learning. Infrared Phys Technol. 2023;129.
96. Sridhar K, Yeruva AR, Renjith PN, Dixit A, Jamshed A, Rastogi R. Enhanced Machine learning algorithms Lightweight Ensemble Classification of Normal versus Leukemic Cells. J Pharm Negat Results. 2022;13:496-505.
97. Gholami E, Tabbakh SRK, Kheirabadi M. Diagnosis of Gastric Cancer via Classification of the Tongue Images using Deep Convolutional Networks. J Inf Syst Telecommun. 2021;9(35):191-6.
98. Gite S, Mishra A, Kotecha K. Enhanced lung image segmentation using deep learning. Neural Comput. 2022.35:22839–22853.
99. Tashtoush Y, Obeidat R, Al-Shorman A, Darwish O, Al-Ramahi M, Darweesh D. Enhanced convolutional neural network for non-small cell lung cancer classification. Int J Electr Comput Eng. 2023;13(1):1024-38.
100. Vasile CM, Udristoiu AL, Ghenea AE, Popescu M, Gheonea C, Niculescu CE, et al. Intelligent Diagnosis of Thyroid Ultrasound Imaging Using an Ensemble of Deep Learning Methods. Medicina-Lithuania. 2021;57(4).
101. Wang WY, Zhou HF, Yan YX, Cheng X, Yang P, Gan LZ, et al. An automatic extraction method on medical feature points based on PointNet plus plus for robot-assisted knee arthroplasty. Int J Med Robot Comput Assist Surg. 2023;19(1).
102. Wongsirichot T, Hanskunatai A. A multi-layer hybrid machine learning model for automatic sleep stage classification. Biomed Eng Appl Basis Commun. 2018;30(6).
103. Yi L, Zhang L, Xu X, Guo J. Multi-label softmax networks for pulmonary nodule classification using unbalanced and dependent categories. IEEE Trans Med Imaging. 2022;Pp.
104. Zang JB, Wang JL, Zhang ZD, Zheng YQ, Xue CY. Improving ECG Classification Performance by Using an Optimized One-Dimensional Residual Network Model. Appl Sci. 2022;12(24).
105. Zarikas V, Papageorgiou E, Regner P. Bayesian network construction using a fuzzy rule based approach for medical decision support. Expert Syst. 2015;32(3):344-69.
106. Zech J, Pain M, Titano J, Badgeley M, Schefflein J, Su A, et al. Natural language-based machine learning models for the annotation of clinical radiology reports. Radiology. 2018;287(2):570-80.
107. Zhang X, Svec M, Tracy R, Ozanich G. Clinical decision support systems with team-based care on type 2 diabetes improvement for Medicaid patients: A quality improvement project. Int J Med Inform. 2021;158:104626.
108. Zhao Q, Che Z, editors. Interpretation research of deep learning ECG classification results based on classification contribution value. Proc SPIE; 2023.
109. Amyar A, Modzelewski R, Vera P, Morard V, Ruan S. Multi-task multi-scale learning for outcome prediction in 3D PET images. Comput Biol Med. 2022;151.
110. Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, et al. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial. Comput Biol Med. 2020;124.
111. Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med Inform Decis Mak. 2020;20(1).
112. Bailoor S, Seo JH, Schena S, Mittal R. Detecting Aortic Valve Anomaly From Induced Murmurs: Insights From Computational Hemodynamic Models. Front Physiol. 2021;12.
113. Jia SQ, Wang Y, Wang WZ, Zhang Q, Zhang X. Value of medical imaging artificial intelligence in the diagnosis and treatment of new coronavirus pneumonia. Expert Syst. 2022;39(3).
114. Narain R, Saxena S, Goyal AK. Cardiovascular risk prediction: A comparative study of framingham and quantum neural network based approach. Patient Prefer Adherence. 2016;10:1259-70.
115. Negro R, Rucco M, Creanza A, Mormile A, Limone PP, Garberoglio R, et al. Machine Learning Prediction of Radiofrequency Thermal Ablation Efficacy: A New Option to Optimize Thyroid Nodule Selection. Eur Thyroid J. 2020;9(4):205-12.
116. Orlando N, Gillies DJ, Gyacskov I, Romagnoli C, D'Souza D, Fenster A. Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. Med Phys. 2020;47(6):2413-26.
117. Qidwai U, Qidwai U, Raja M, Burton B. Smart AMD prognosis through cellphone: an innovative localized AI-based prediction system for anti-VEGF treatment prognosis in nonagenarians and centenarians. Int. Ophthalmol. 2022;42(6):1749-62.
118. Rafael-Palou X, Aubanell A, Ceresa M, Ribas V, Piella G, Ballester MAG. Prediction of Lung Nodule Progression with an Uncertainty-Aware Hierarchical Probabilistic Network. Diagnostics. 2022;12(11).
119. Göltepe Y. Performance of lung cancer prediction methods using different classification algorithms. Comput Mater Continua. 2021;67(2):2015-28.
120. Han X, Hong J, Reyngold M, Crane C, Cuaron J, Hajj C, et al. Deep-learning-based image registration and automatic segmentation of organs-at-risk in cone-beam CT scans from high-dose radiation treatment of pancreatic cancer. Med Phys. 2021;48(6):3084-95.
121. Hasimbegovic E, Papp L, Grahovac M, Krajnc D, Poschner T, Hasan W, et al. A Sneak-Peek into the Physician's Brain: A Retrospective Machine Learning-Driven Investigation of Decision-Making in TAVR versus SAVR for Young High-Risk Patients with Severe Symptomatic Aortic Stenosis. J Pers Med. 2021;11(11).
122. Hossain MM, Hasan MM, Rahim MA, Rahman MM, Yousuf MA, Al-Ashhab S, et al. Particle Swarm Optimized Fuzzy CNN With Quantitative Feature Fusion for Ultrasound Image Quality Identification. IEEE J Transl Eng Health Med. 2022;10:1800712.
123. Jung JW, Hwang S, Ko S, Jo C, Park HY, Han HS, et al. A machine-learning model to predict postoperative delirium following knee

arthroplasty using electronic health records. BMC Psychiatry. 2022;22(1).
124. Işik AH, Güler I, Şener MU. A low-cost mobile adaptive tracking system for chronic pulmonary patients in home environment. Telemed J E Health. 2013;19(1):24-30.
125. Sumana G, Kalaiselvi K, Vijayalakshmi J, Shanthi S, Aparna G, Joseph MK. A design and development of support system for prediction of various renal syndromes using artificial neural networks. Int J Syst Assur Eng Manag.
126. Tang J, Fang Y, Xu Z. Establishment of prognostic models of adrenocortical carcinoma using machine learning and big data. Front Surg. 2023;9.
127. Vodenčarevic A, Van Der Goes MC, Medina OAG, De Groot MCH, Haitjema S, Van Solinge WW, et al., editors. Predicting flare probability in rheumatoid arthritis using machine learning methods. DATA 2018 - Proceedings of the 7th International Conference on Data Science, Technology and Applications; 2018.
128. Werneburg GT, Werneburg EA, Goldman HB, Mullhaupt AP, Vasavada SP. Machine learning provides an accurate prognostication model for refractory overactive bladder treatment response and is noninferior to human experts. Neurourol Urodyn. 2022;41(3):813-9.
129. Wu CH, Chiu RK, Wang SA. A cloud-based fuzzy expert system for the risk assessment of chronic kidney diseases. Int J Bus Syst Res. 2015;9(4):315-33.
130. Zhang IY, Hart GR, Qin B, Deng J. Long-term survival and second malignant tumor prediction in pediatric, adolescent, and young adult cancer survivors using Random Survival Forests: a SEER analysis. Sci Rep. 2023;13(1).
131. Zimmermann CJ, Zelenski AB, Buffington A, Baggett ND, Tucholka JL, Weis HB, et al. Best case/worst case for the trauma ICU: Development and pilot testing of a communication tool for older adults with traumatic injury. J Trauma Acute Care Surg. 2021;91(3):542-51.
132. Vallée A, Sveltlane D, Trichereau J, Neveu S, Fourn E, Majerholc C, et al. Electronic medical record alert increases HIV screening rates: the Foch hospital pilot POP-up project. BMC Health Serv Res. 2022;22(1).