

Received July 28, 2020, accepted August 11, 2020, date of publication August 14, 2020, date of current version August 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3016782

Implementation and Use of Disease Diagnosis Systems for Electronic Medical Records Based on Machine Learning: A Complete Review

JAHANZAIB LATIF¹, CHUANGBAI XIAO¹, SHANSHAN TU¹, (Member, IEEE),
SADAQAT UR REHMAN², (Member, IEEE), AZHAR IMRAN³, (Graduate Student Member, IEEE),
AND ANAS BILAL¹

¹Engineering Research Center of Intelligent Perception and Autonomous Control, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

²Department of Computer Science, Namal Institute, Mianwali 42250, Pakistan

³School of Software Engineering, Beijing University of Technology, Beijing 100124, China

Corresponding author: Chuangbai Xiao (cbxiao@bjut.edu.cn)

This work was supported in part by the China National Key Research and Development Program under Grant 2018YFB0803600, in part by the Natural Science Foundation of China under Grant 61701009 and Grant 61801008, in part by the Beijing Science and Technology Planning Program of China under Grant Z171100004717001, in part by the Beijing Natural Science Foundation under Grant 4172002 and Grant L172049, in part by the Scientific Research Common Program of Beijing Municipal Commission of Education under Grant KM201910005025, and in part by the Chinese Postdoctoral Science Foundation under Grant 2020M670074.

ABSTRACT Electronic health records are used to extract patient's information instantly and remotely, which can help to keep track of patients' due dates for checkups, immunizations, and to monitor health performance. The Health Insurance Portability and Accountability Act (HIPAA) in the USA protects the patient data confidentiality, but it can be used if data is re-identified using 'HIPAA Safe Harbor' technique. Usually, this re-identification is performed manually, which is very laborious and time captivating exertion. Various techniques have been proposed for automatic extraction of useful information, and accurate diagnosis of diseases. Most of these methods are based on Machine Learning and Deep Learning Methods, while the auxiliary diagnosis is performed using Rule-based methods. This review focuses on recently published papers, which are categorized into Rule-Based Methods, Machine Learning (ML) Methods, and Deep Learning (DL) Methods. Particularly, ML methods are further categorized into Support Vector Machine Methods (SVM), Bayes Methods, and Decision Tree Methods (DT). DL methods are decomposed into Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Deep Belief Network (DBN) and Autoencoders (AE) methods. The objective of this survey paper is to highlight both the strong and weak points of various proposed techniques in the disease diagnosis. Moreover, we present advantage, disadvantage, focused disease, dataset employed, and publication year of each category.

INDEX TERMS Automatic extraction, classification, clinical informatics, deep learning, disease diagnosis, electronic health records, machine learning.

I. INTRODUCTION

The term diagnosis is used for finding symptoms of disease or analysis of the patients to determine the health conditions. The diagnosis is usually performed through one of these methods, i.e., examining the physical condition of the patient, exploring patient's history, or from diagnostic tests which are analysed by various healthcare professionals such as den-

tist, physician, chiropractor, physical therapist, or physician assistant and compounder etc. [1]. The patient's history is frequently saved in the form of a prescription for necessary medications, streamline workflow, and to keep track of the patient's performance. Initially, the prescription was saved in the form of the paper chart containing the type of diseases, suggested medicines, vaccination dates, treatment plans, and the test results of X-rays specific hospitals. However, in the modern age of the computer, the prescription is saved in a digital format which is known as an electronic

The associate editor coordinating the review of this manuscript and approving it for publication was Xiong Luo¹.

medical record (EMR) or electronic health record (EHR). These electronic records help the physicians to access the patient's records instantly, to keep track of patients' due dates for checkups and immunisations and monitor patients health performance and make decisions accordingly [2].

Although both these terms (EMR and EHR) are used interchangeably, according to the 'Office of the National Coordinator of Health Information Technology (ONC)' both the terms are utilised exclusively [3]. The EMR is the digital form of the prescription, which contains the patients' information collected in a provider's office for healthcare professionals. The EMR data can be either human generated or machine generated [4]. EMRs have multiple advantages over paper prescriptions including instant access, keep track of patients' information, save patients visits, screen patients, and enhance the healthcare's quality [5]. The scope of the EHRs is going beyond than EMRs, as it contains information of all the medical investigator involved in patients' health records. The patient's information is also shared with other clinicians and medical researchers in various hospitals to study and improve the root causes of the disease. EMRs contain temporal and heterogeneous doctor order information which may used as an input for treatment pattern discovery [6].

EHRs also facilitate the patients to see their records on how the progress is going on, which motivates them in many cases (not necessarily) [7], [8]. Though EMR and EHR provides many benefits to users and the practitioners, but there are a lot of challenges associated with the implementation of these electronic records including downtime of the computers, incapability of the computer experts, lack of communication among users and, security threats of confidentiality leakage etc. [9]–[11]. Though, these methods are the decent source of data collection in various hospitals to assist clinicians for the manual extraction of the relevant information. But, still, it is very tedious and time taking exertion to extract specific information from the bulk of data. There is a need for automatic systems for disease diagnosis from electronic medical records. Currently, EHR systems have been widely adopted by different healthcare professionals and institutions to provide fast, efficient and real-time healthcare services in economical ways [12], [13].

There are various healthcare systems has been developed to overcome the challenges of electronic records, including the volume of EHRs data, incompleteness, and inconsistency of the data. NLP based computational phenotyping has various applications such as new phenotype identification, diagnosis classification, and clinical trial screening by implementing different methods, i.e. rule-based methods, machine learning methods, and deep learning methods [14]. They are employed to transform the raw text into useful information, also known as case-base information in 'Case-Based Reasoning (CBR) Systems' [15].

We present the survey of various techniques employed for automatic disease diagnosis from the electronic record in this paper. The main aim of this survey is to provide a platform to researchers to review different disease diagnosis systems;

to provide a detailed description of their work; to highlight advantages and disadvantages of the various approaches; to discuss present trends and the future directions, and to give a comparative analysis of the multiple techniques. The detail description of this review is given in Fig.1. The number of articles reviewed with respect to publication year and concerned disease is given in Fig.2 and Fig. 3. respectively.

The paper is structured into five sections. The existing work on the disease diagnosis approaches is discussed in Section 2. In Section 3, the challenges associated with these approaches are addressed. The discussion is given in section 4. Finally, the conclusion of the survey and future trends are presented in Section 5.

II. DATA-DRIVEN MODELS FOR DISEASE DIAGNOSIS

The researchers have designed multiple techniques for computer-aided diagnosis of diseases. This diagnosis has assisted in clinical and biomedical research by employing many applications, i.e., phenotype discovery, diagnosis classification, and patient screening etc. We have categorised various disease diagnosis methods based on their methodology. The disease diagnostic tasks are classified into rule-based methods, support vector machine methods, neural network methods, Bayesian methods, and deep learning methods.

A. RULE-BASED METHODS

Rule-based methods are used in computer science to learn and identifies rules and apply to extract specific patterns associated with that rule [16], [17]. These rules are used to identify and utilize a set of relational rules for capturing the knowledge, unlike the machine learning systems in which a single model is applied for the task of prediction. The rule-based methods include rule association mining [18], learning classifier systems [19], artificial immune systems [20], cloud based model using Fuzzy logic concepts [21], and other techniques based on rules. These methods are widely adopted by the researchers for phenotype cohort identification of the patients. The review of Shivade *et al.* [22] contains 24 rule-based methods out of 97 methods.

In [23], a rule-based method has been used to identify Cardiac Resynchronization Therapy from clinical records. A dataset of Clinical Translation Science Institute of University of Minnesota's consists of 6174 reports (training set 3700, test set 2474 reports) was used to match with New York Heart Association custom diagnosis codes designed locally by the hospital. The proposed method is compared with SVM method. The results of this study have shown that the rule-based method performed similarly like machine learning methods. The author has achieved state of the art results with precision, recall, and F-measure scores 94.99%, 92.13%, and 93.37%, respectively. This work can be advantageous by performing more validation of data and results.

A rule-based method using forward chaining in electronic health record has been proposed in [24]. The author has used the codes of International Classification of Disease (ICD-10) for the patients Haemogram reports. The forward

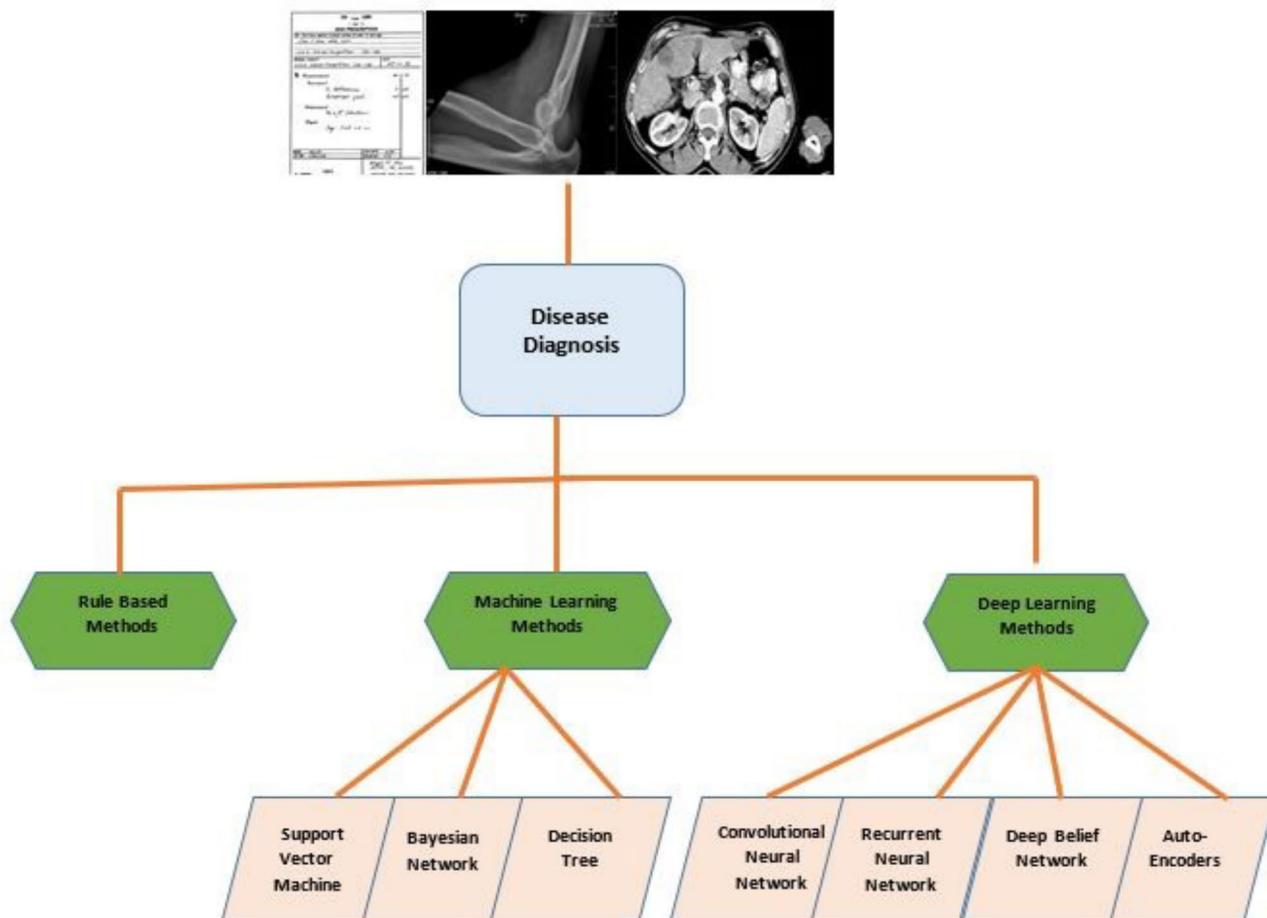


FIGURE 1. Breakdown of the disease diagnosis methods in this survey.

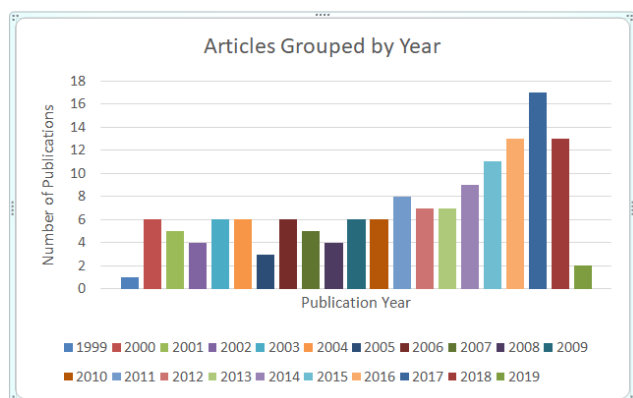


FIGURE 2. Articles grouped by year.

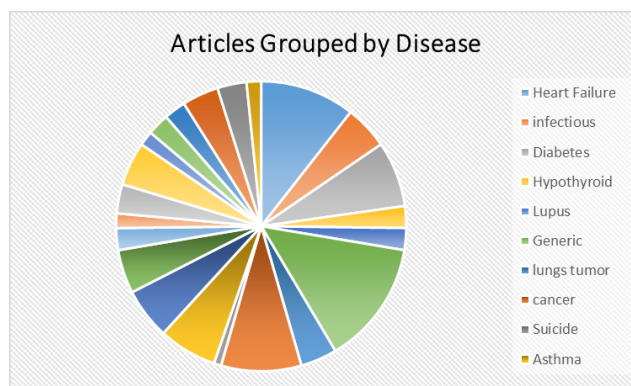


FIGURE 3. Articles grouped by concerned disease.

chaining (FC) rule-based method has been utilised to extract healthcare indicators in the dataset. The proposed method has been extracted information from patient records based on the ICD-10 code alike the method proposed in [25], which successfully extracted information with average f-measure of 96.83%.

Jorge *et al.* [26] have been employed rule-based methods to identify lupus patients from EHR dataset with 400 records

(200 each for validation and training). The narrative and codified information have been extracted using NLP from the training set data. The author has been designed algorithm using ‘penalised logistic regression’ to categorise definite Systemic Lupus Erythematosus (SLE) and definite/probable SLE. The results of the top-performing rule-based algorithm (ICD-9 Code) has shown (86%, 84%) specificity,

(60%, 69%) sensitivity for definite SLE and definite/probable SLE respectively.

A rule-based approach has been employed in [27] for the detection of adverse drug events (ADE). A dataset composed of 115447 records of Danish, French and Bulgarian hospitals. Association rules and decision tree methods are used to discover ADE detection rules concerning time constraints. The field experts have been manually validated, filtered, and reorganised the rules which were then placed into rules repository. There are 236 validated ADE detection rules have been discovered which were able to detect 27 various outcomes. The rules associated with anticoagulant drugs (35%), hyperkalemia (27%), and pharmacokinetic drug (25%) have been discovered automatically in this study.

Detailed medical information and relationship between text instances were extracted as a co-reference resolution from a patient record summary in [28]. Three NLP systems employed to resolve phrased which refer to the same entity, i.e. rule-based model, a maximum entropy model, and Markov logic network model (MLN). The results of this study have shown that the proposed model has achieved 4.3% and 5.7% F-scores on Beth and Partners datasets respectively. These three models are integrated into an ensemble system to generate state of the art results as compared to baseline systems with 87.21% performance measure which is 4.5% higher than i2b2 Track 1C average.

In [29], temporal trends have been discovered from the Nationwide Inpatient Sample (NIS) dataset. A combined method of rule mining and model-based recursive partitioning was used to discover temporal trends as a whole, but also for significant subgroups (based on patient age, gender, disease). The results of this study have proved that the trends discovery of subset based on the age-sex relationship is impossible using trend-tracking methods. The dataset consists of only 20% of the discharged patients in the USA; the significance of this research can be improved by adding more number of patient's data.

Petersen *et al.* [30] have been identified patients with coronary artery bypass or cardiac catheterisation using ICD-9. The dataset composed of 5151 patient records of Veteran Health Administration has been utilized in this study. The rule based method to diagnose artery disease, childhood obesity, and peripheral artery disease has achieved above 95% of sensitivity and specificity score. The positive predictive value of acute myocardial infarction is better than existing methods.

In [31], McCoy *et al.* have been performed a comparative analysis of association rule mining and crowdsourcing techniques for proper information extraction from electronic health records. These methods are evaluated on publically available EHR in which rule mining and crowdsourcing approach recognised 19586 and 31440 pairs respectively. The author has compared only 500 pairs of each in which 186 pairs were overlapping. The results of this study have shown that crowdsourcing approach has been identified most common relationships unlike association rule mining, which recognized only rare relationships. The better

results can be obtained if both these methods are used in combination.

A rule-based system has been designed by Wiley *et al.* for the detection of stain induced myotoxicity [32]. The manual annotation has been performed of 300 allergy patients, and keywords are defined for this dataset. The set of rules have been used to detect contextual information specific to the identified keywords. They achieved 86% and 91% scores of positive predictive value (PPV), and negative predictive value (NPV), respectively.

In [33], the author has been developed a list containing concepts and secondary concepts found in the same sentence. The secondary concepts mainly contained medications. A set of rules have been defined after concepts initialization for the identification of phenotypes. The proposed method achieved 92% of the kappa value along with the original annotations. A similar concept based technique has also been implemented in [34] which achieved 0.72, 0.78, and 0.94 T, N, and M staging accuracies.

A rule-based heuristic approach has been used in [35], for the assertion of colorectal cancer. The colorectal-cancer concepts detected using MedLee approach [36]. They searched the concept contexts by applying a predefined set of rules. The proposed method achieved 99.6% F-measure for document level concept identification.

In [37], a rule-based approach has been developed by Li *et al.* to detect adverse drug events (ADE) and medical errors as in [27]. The author detected ADE from patient health records, clinical tests, and medications. The performance of the proposed method is evaluated on a trigger tool [38], and they reached a 100% agreement. The triggers are generally a combination of keywords that are used to extract specific information of the underlying disease.

Alodadi [39], proposed a system to extract radiology notes from electronic health records and discovered association rules to identify interrelationship between various medical entities. A textual dataset has been used to generate transactional records of each radiology report. The Bag-of-Concepts used to formalise raw clinical text and concept's weight was measured using text mining weight methods, i.e. TF-IDF (Text Frequency-Inverse Document Frequency). The weight of every item introduced using Concept Unique Identifiers (CUIs). The proposed methods have shown significant results because of three main measures, i.e. confident, lift, and weighted support.

Szenasi *et al.* [40] developed a rule-based method for concept extraction from unstructured medical records. NLP techniques employed to process medical text. The text has been decomposed into sentences in which token was identified. A Systematized Nomenclature of Medicine - Clinical Term (SNOMED-CT) was used to make queries for the identification of concepts. Three filters have been applied to obtain better results of the proposed method, i.e. hierarchy based filter, POS filter, context similarity filter, and hierarchy filter which are used in a sequence to reduce computational complexity. A Medline dataset was used to evaluate the method,

which has achieved 88.77%, and 89.69% recall and precision respectively.

In [41], a rule-based grammar approach has been used to extract textual information from mamma carcinoma patient's records. The therapy suggestion was derived by defining seven significant variables from the extracted textual features. The proposed system was evaluated on the mammography use case. An integration technique based on rule-based decision support, semantic modelling and information extraction has been used to extract textual feature with 0.69 of accuracy and achieved 0.90 for lymph node status.

A novel algorithm has been developed to identify suicide, or suicidal ideation in electronic health records [42]. ICD-9 was applied to collect various patient's records from 2004 to 2010, in collaboration with the Food and Drug Administration. A training set of 50 records are used, classified as positive and negative by the experts. The study has shown that ICD version 9 had 0.55 positive predictive value (PPV), but ICD-09 in combination with NLP had PPV value of 0.97.

Sauer *et al.* have been used the rule-based approach on structured and semi-structured Veteran affairs data to extract pulmonary function test (PFT) records [43]. They used the NLP tool to retrieve Spirometric values and responses to bronchodilator challenge. A random set of 1001 documents are used to evaluate the performance of the model which achieved 98.9%, 98.8%, and 98.3% of precision, recall, and f-measure, respectively.

B. MACHINE LEARNING METHODS

Machine learning (ML) methods are most commonly used to construct medica database system from electronic health records for those patients who have undergone health examination [44]. These ML methods are subdivided into support vector machine methods, bayesian methods, and decision tree methods.

1) SUPPORT VECTOR MACHINE METHODS

Support Vector Machine (SVM) is one of the commonly used algorithms adopted by researchers for supervised classification. Vapnik developed SVM in 1990, which works by labelled data [45]. When both input and output are already given, then the input is used as a training set to classify the data as like output.

In [46], cancer diagnosis has performed on medical records extracted from EHR using SVM model. The proposed model was evaluated on 100 and 400 pieces of medical records for 10 and 3 different types of cancers and achieved 86.2% and 97.33% of prediction accuracy, respectively. The Radial Basis Function (RBF) kernel along with K-fold cross-validation techniques was used to estimate the performance of the model. The higher prediction accuracy may be achieved by evaluating the proposed model on a larger dataset.

Brisimi *et al.* [47] has been focused on various diseases, e.g. chronic, heart and diabetes from electronic health records and considered as a binary classification problem. They employed various machine learning methods including

Support Vector Machine (sparse and kernel), logistic regression (LR) and random forest (RF) for this binary classification problem. These methods are evaluated on the dataset of Boston Medical Center, England. They majorly contributed by using two novel methods K-LRT (likelihood ratio) and JCC (joint clustering and classification) methods. JCC has been applied to find out hidden clusters in positive samples and to recognize sparse classifiers of a cluster by separating positive and negative samples (hospitalized vs non-hospitalized).

In [48] SVM has been applied as a hybrid model by combining the features of weighted LDA and word vector model for liver disease classification. The hybrid model has been applied to get standard features, which were further utilized on the classification to get optimum results. The proposed method evaluated on cooperative hospital dataset (Sogou, Baidu, and Tencent) and achieved 99.1% accuracy over 100 dimensions. This combination of weighted LDA and word vector model gives the advantage of improved prediction accuracy, diagnostic reliability and full use of possible text information of EHR data.

Alemzadeh and Devarakonda [49] used SVM, NB, DT along with ensemble methods AdaBoost and Random Forests for automatic identification of disease control status from EHR data. The dataset used in this research was comprised of 55,000 clinical notes over a period of seven years (average of 133 notes per patient). The system has been evaluated on 5035 candidate snippets (2086 labels, 2949 unlabeled) and achieved 86% F-measure score in identifying disease status and 77% accuracy in classifying the status is controlled or not.

A machine learning method SVM has been applied on naive and expert-defined EHR features collection in [50] to detect Rheumatoid Arthritis (RA) cases with the help of medication exposures, billing codes, and NLP concepts. The Training of the SVM method has been performed on both naive and expert-defined data which had achieved precision and recall scores 0.94 and 0.87 respectively unlike the deterministic algorithms with 0.75 and 0.51 of precision and recall. A dataset of 10,000 patients was used in this research which was further classified as possible RA, definite RA, and not RA based on the test results. The validity of this method can be verified on different diseases.

A component-based NLP system HITEx has been developed by Zeng *et al.* [51] to extract findings for disease research. A record of 50 patients was used in the HITEx system to extract smoking, co-morbidity, and principal diagnosis. The SVM classifier was used to identify the smoking status of patients. It was tested and trained on 8500 smoke sentences with 10-folded cross evaluation and experimented on Weka tool. The HITEx achieved an accuracy of 0.82, 0.87, and 0.90 on principal diagnosis, co-morbidity, and smoking status extraction, respectively excluding insufficient data which has outperformed ICD-9. The better results of sensitivity, specificity, and accuracy can be obtained by combining the proposed method with ICD-9.

TABLE 1. Summary of rule-based methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[26]	Jorge et al.	Rule-based algorithm, Machine Learning codified algorithm	2018	400 subjects	Definite SLE	Sensitivity: 86%, Specificity: 60%, PPV: 46%	Identifying Lupus Patients from EHR
					Definite Probable SLE	Sensitivity: 84%, Specificity: 69%, PPV: 65%	
[23]	Zhange et al.	Rule-Based Method, SVM: bag of words, SVM: n-gram	2017	University of Minnesota's 6174 clinical notes	Heart Failure	Precision: 94.99%, Recall: 92.13%, F-Measure: 93.37%	Extract New York heart association classification from EHR
[41]	Breischneider et al.	Rule-based grammar approach	2017	8766 clinical reports of 2096 patients collected from University Hospital Erlangen	Breast Cancer	Accuracy: 90%, Sensitivity: 98%, Specificity: 59%	Automatic extraction of breast cancer information from clinical reports
[32]	Daymont et al.	Simulated error identification	2017	2,000,595 growth measurements of 280,610 patients	---	Sensitivity: 97%, Specificity: 90%	Automated identification of implausible values in growth data from pediatric EHR
[43]	Sauer et al.	NLP	2016	Rocky Mountain Network dataset with 1001 documents	Asthma	Precision: 98.9%, Recall: 97.8%, F-measure: 98.3%, Accuracy: 99.7%	Extract PFT Reports from Structured and Semi-structured Veteran Affairs Data
[25]	Lakshmanan & Arumugam	ARBIE	2015	UCI dataset with 270, 426, and 1172 records for heart attack, diabetes and hypothyroid	Heart Attack	Precision: 96.12%, Recall: 96.88%, F-measure: 96.49%	Extract information from the patient's data using Advance Rule-Based Information Extraction
					Diabetes	Precision: 96.59%, Recall: 96.12%, F-measure: 96.35%	
					Hypothyroid	Precision: 97.81%, Recall: 97.48%, F-measure: 97.64%	

TABLE 1. (Continued.) Summary of rule-based methods.

[40]	Szenasi et al.	Ontology-based NLP	2015	Medline abstracts dataset USA	Multiple Diseases	Recall: 88.77%, Precision: 89.69%	Concept extraction from contextual approach by querying the SNOMED-CT medical-ontology and three filtering rules
[?]	Kumar & Padmapriya	Forward Chaining	2014	500 records of patient medical examination reports	Multiple diseases including Cholera, and Typhoid	---	Extracted information from a patient medical report based on ICD-10 code using Forward-Chaining Rule System
[29]	Hrovat et al.	Temporal trend discovery	2014	Nationwide Inpatient Sample dataset with 65,308,185 records	Ten diseases including, hypertension, diabetes, and heart failure etc.	---	Contrasting temporal trend discovery for large healthcare database
[37]	Li et al.	AEME detection	2014	753 NICU patients	AE and ME detection	PPV: 100%, Sensitivity: 100%, Specificity: 100%	AE or ME detection using EHR in NICU in the USA
[28]	Dai et al.	Multi-stage rule-based (MR) system, Markov Logic Network-based system, Maximum Entropy-based Co-reference Resolution System	2012	Three datasets including Beth, Partners, and University of Pittsburgh dataset	Multiple Diseases	F-score (Beth): 86.6%	Co-reference resolution of medical concepts in discharge summaries by exploiting contextual information
						F-score (Partners): 87.4%	
[31]	McCoy & Wright	Association rule mining and Crowdsourcing	2012	418,221 medications for 53,108 patients	Multiple Diseases	---	Comparison of Association Rule Mining and Crowdsourcing for Automated Generation of a Problem-Medication Knowledge Base
[42]	Haerian et al.	NLP	2012	280 patients EHR records	Suicide or Suicidal Ideation	PPV: 90%, Kappa: 86.7%	Identifying Suicide or Suicidal Ideation by combining ICD-9 and NLP from EHRs

TABLE 1. (Continued.) Summary of rule-based methods.

[27]	Chazard et al.	Data mining based induction	2011	115,447 stays from French, Danish, and Bulgarian hospitals	ADE detection Hyperkalemia	Recall: 95.1%, Precision: 52%, F-Measure: 67.2%	Data Mining to Generate ADE Detection Rules
[35]	Xu et al.	Rule-based, ML-based	2011	Synthetic Derivative database of 1262671 patients	Colorectal Cancer	F-measure: 99.6%, Precision: 99.6%, Sensitivity: 99.6%, Specificity: 96.9%, Accuracy: 99.6%	Extracting and Integrating Data from EHR for Detecting Colorectal Cancer
[34]	Nguyen et al.	Rule-based SNOMED-CT	2010	718 lung cancer patients	lung tumor-node-metastases cancer	Accuracy T: 72%, N: 78%, M: 94%	Rule-based MEDTEX system using SNOMED CT ontology used to extract lung TNM staging factors
[33]	Ware et al.	NLP	2009	700 discharge summaries	16 diseases related to obesity including Depression, Asthma, and Diabetes	Kappa values Textual 92% and Intuited 91%	An NLP framework to assess clinical conditions
[30]	Petersen et al.	Positive predictive value	1999	5151 discharge summaries of Veteran Health Administration	Myocardial Infarction	PPV: 96.9%, Sensitivity: 96%, Specificity: 100%	Positive predictive value of the diagnosis of MI

An NLP system has been designed in [52], [53] for pneumonia identification from narrative reports of the patient. This system was considered as a binary classification problem to predict whether a patient was positive for pneumonia or not. A dataset of 426 patients (unrestricted) was employed in this research during their stay in ICU. There were various features considered in research, i.e. unified modelling language system, word n-gram, and assertion value for pneumonia. The performance of the system was improved by utilizing feature selection approach characterized by statistical testing for pneumonia identification. A comparative analysis was performed on a restricted dataset of 236 patients. The proposed method has shown significant results of F-measure (0.8571 and 0.8176) and (0.5070 and 0.4910) on the restricted and unrestricted datasets, respectively.

Garla *et al.* [54] compared Laplacian and linear SVM for classification task based on clinical text along with the effect of semi-supervised learning were determined on Laplacian SVM performance. The training of SVM and Laplacian

SVM was performed on 820 reports of ultrasound, MRI, and abdominal CT labelled for malignant liver lesions (77% positive class presence). Furthermore, they have used 19845 unlabeled samples (random) along with gold standard references for Laplacian SVM. A test set of 520 labelled reports was used for both these methods. The results of Laplacian SVM (labeled and unlabeled) has been outperformed supervised SVM (0.943 vs 0.911 sensitivity, 0.877 vs 0.883 PPV, and 0.773 vs 0.741 F1-score). The semi-supervised methods like Laplacian SVM are implemented on different unlabeled EHR data to improved clinical text classification.

An SVM based method has been employed to identify contralateral breast cancer from narrative text and pathology reports [55]. The medical concepts and their combinations are used in the proposed method for the identification of contralateral events in the notes. Moreover, pathology reports of the breast (either left or right) are used as an additional feature. SVM along with derived features are used to detect contralateral cancer. They achieved 0.93 and 0.89 area under

TABLE 2. Summary of SVM methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[46]	Zhang et al.	SVM-RBF	2019	100 and 400 medical records per cancer from New York Clinics	Cancer	Accuracy: 97.33%	Cancer classification from EHR using SVM
[47]	Brisimi et al.	K-LRT and JCC	2018	45579 heart and 33122 diabetes patients of Boston Medical Center, England	Heart, Diabetes	Accuracy- H: 77.06%, D: 78.55%	Predicting Chronic Disease from EHR using K-LRT and JCC
[48]	Jin et al.	LDA with Word-Vector Model	2017	Cooperative hospital dataset of Sogou, Baidu, and Tencent	Liver	Accuracy: 99.1%	Liver disease diagnosis from EHR using LDA and word vector model
[49]	Alemzadeh et al.	NLP-based cognitive system	2017	55,000 clinical notes of 414 patients collected from US hospital	Hypertension	F1-score: 86%, Accuracy: 77%	Disease status detection in EHR by employing NLP
[55]	Zeng et al.	SVM	2017	1063 breast cancer patients	Breast Cancer	AUC: Validation: 93%, Testing: 89%	Identification of contralateral breast cancer using extracted features and pathology reports.
[54]	Garla et al.	Laplacian SVMs	2013	1340 (819 training and 521 testing) abdominal CT, MRI reports	Cancer	Macro-F1: 77.3%, Sensitivity: 94.3%, PPV: 88.3%, Kappa: 54.8%, F1: 91.2%, Specificity: 55.7%	Clinical text classification for Cancer with Laplacian SVMs
[56]	Chen et al.	Active learning with SVM	2013	Three datasets of RA (376), CRC (300), and VTE (703)	RA, CRC, and VTE	AUC: 95%	Applied active learning with SVM for RA, and colorectal cancer
[52]	Bejan et al.	statistical feature selection	2012	426 and 236 patients (unrestricted and restricted dataset)	Pneumonia	F1-measure Unrestricted: 85.71%, Restricted: 81.67%	Pneumonia identification using statistical feature selection
[50]	Carroll et al.	SVM	2011	376 patients population (185 RA, 191 not RA)	Rheumatoid Arthritis	Precision: 94%, Recall: 87%, F-Measure: 88.6%, AUC: 96.6	%Naive EHR Phenotype Identification for RA using SVM.
[57]	Wei et al.	SVM	2010	Mayo clinic's medical data of 1600 T2DM patients	Diabetes	Precision: 96.8%, Recall: 94.3%, F-Score: 95.6%	Patient identification from T2DM using ML
[51]	Zeng et al.	SVM, Naive Bayes, and Decision Tree	2006	150 discharge summaries	Asthma	Accuracy (PD): 82%, Accuracy (CM): 87%, Accuracy (SS): 90%	Extracting PD, CM,SS for asthma using NLP

the curve for the validation and test set, respectively. This method can be used to detect various disease because of the feature generation simplicity, and it can also be applied to other breast cancer events like distance or local recurrence.

A novel method has been proposed in [57] by combining machine learning, NLP, and ontology methods for patient identification. A dataset of Mayo Clinic electronic records from 2007 has been used to evaluate this method. SVM algorithm experimented on the extracted T2DM (type 2 diabetes mellitus) clinical notes of SNOMED concepts. The proposed methods are evaluated on precision, recall, and F-measure which achieved 95% f-measure score. The evaluation of T2DM is a minimal area; the proposed method can be applied to other diseases to attain better performance. A sparse discriminant analysis (SDA) has been used to predict heart disease patients using electronic health records. SDA was evaluated on 280 instances of heart disease dataset, and achieved 96% prediction accuracy. The proposed method was preferred over linear discriminant analysis (LDA) because of linear inseparability and poor evaluation of LDA parameters [58].

Chen *et al.* [56] have been used active learning (AL) with a support vector machine (SVM) based method for electronic health data. The performance of the proposed method is evaluated on three diseases cohorts' i.e. colorectal cancer (CRC), venous thromboembolism (VTE), and rheumatoid arthritis (RA) using two feature sets, i.e. unrefined features (all clinical concepts & billing codes) and refined features (domain experts selected). A performance comparison was made between active and passive learning (PL) based on random sampling. The results of AL outperformed PL on three phenotype tasks. AL reduced 68% and 23% of the annotated samples respectively when the unrefined feature was employed for RA and CRC and achieved 0.95 area under the curve (AUC). Furthermore, AL reduced 68% annotated samples for VTE with 0.70 AUC using refined features. The performance of the phenotype classifiers is improved using refined features. An efficient phenotype method can be developed by combining AL and feature engineering characterized by domain knowledge.

2) BAYESIAN METHODS

The Bayesian Network (BN) and the Naive Bayes (NB) are probabilistic algorithms and both work elegantly with a large number of features [51]. There is no need for a dependency network for NB classifiers and works better with high dimensional features, unlike Bayesian network. In Naive Bayes, patterns are matched by examining a set of categorized documents. It is a probabilistic classifier that matches the data with the bag of words. The major aim of this classifier is to classify different documents into certain class or category. It streamlines the learning by classifying the features in an independent class. Naive-Bays accuracy is independent of the feature's dependencies in the class [59].

A Naive Bayes classification has been used to predict heart disease from electronic health records in [60]. The proposed system used to history based heart disease dataset to extract hidden knowledge (heart disease). This system utilized continuous data rather than classified data which has achieved significant results. The future works include an increase in data volume and by focusing on other diseases like liver and cancer to obtain better results.

Al-Aidaros *et al.* [61] presented a Naive Bayes method for medical data classification. The classification of various diseases like breast cancer, liver disorders, lung cancer, and primary tumour etc. based on accuracy and area under the ROC curve has been performed. The proposed method was compared with five other methods evaluated on 15 datasets. The results of this have shown that NB outperformed others in medical classification. The proposed method followed by deep learning concepts can be employed to obtain better segmentation results. The future works include hybridization of NB with other approaches.

The Bayesian and decision tree methods of medical diagnosis have been initialized in [62]. These methods were employed to identify diagnostic problems like a diagnosis of thyroid diseases, primary tumour localization, rheumatology, and analytical reappearance of breast cancer. The NB has shown promising accuracy. The redundant knowledge can be used with NB for better performance and to deal with missing values.

Wang *et al.* [63] have been employed a Bayesian approach to predict the occurrence of brain metastasis from lung cancer. A global dataset comprised of 50 thousand cancer records obtained over the period of 1996 to 2010 in Taiwan was used in this research. The proposed has shown state of the art results in terms of sensitivity compared with SVM, and LR evaluated over performance metrics, i.e. accuracy, sensitivity, and specificity.

Bayesian classifier has been used to assess the progression or relapse of cancer in [64]. A brain tumour dataset of 142 patients collected from 2000 to 2005 was used in this study. A total of 96 attributes (binary form) were selected for training. The proposed model has been calculated the probability of being consigned to relapse and no-relapse risk of cancer. The proposed method achieved good results of

accuracy, specificity, and sensitivity with values 0.84, 0.87, and 0.80, respectively.

Sakai *et al.* [65] employed a Bayesian network for diagnostic prediction of appendicitis. A dataset composed of 169 patients suspected of acute appendicitis was considered in this research. The proposed model was compared using performance metrics (error-rate and area under the ROC curve) with logistic regression, and neural network model. This model detected that 86 out of 169 patients were suffering from appendicitis and achieved the lowest error rate and reliable results of ROC curve.

Naive Bayes method has been used to predict heart diseases in [66]. A clinical dataset of 500 patients obtained from the diabetic research institute Chennai was used in this study. The different attributes like age, sex, diagnosis, and slope were selected for the proposed model. The performance of the method was evaluated using performance matrices like precision, recall, and f-measure, and achieved 0.71, 0.74, and 0.712 of these values, respectively.

In [67], genotype data has been to predict Alzheimer's disease using a Bayesian network. The genome data of 1411 patients were collected for this study (860 with Late-Onset Alzheimer Disease (LOAD), and 550 without LOAD). The performance of the proposed method was compared with the model- averaged (MANB), feature selection (FSNB), and without feature selection (NB) in terms of running-time, AUC, and calibration. MANB has been outperformed others based on these performance metrics with training time 16s and 0.72 of AUC to predict LOAD patients.

3) DECISION TREE METHODS

A decision tree is one of the commonly used algorithms for data analysis [68]. It contains terminal and non-terminal nodes. Each non-terminal node depicts a condition or test on a data item. This technique is generally used for the classification of records which is helpful for both association and regression tasks. The instances are sorted in a decision tree from non-terminal to terminal nodes [69]. It is easy to visualize and recognize the various advantages and disadvantages of data using decision tree [70]. There is a lot of work has been done for disease diagnosis using a decision tree.

In [71], the decision tree algorithm (C4.5 decision tree classifier) has been used for the diagnosis prediction of dengue fever in the primary stage of infection. The screening of 1200 patients (1012 from Singapore and 188 from Vietnam) has been carried out during the first 3 days of illness and followed up to the period of one month. A total of 364 patients has been found suffering from dengue (171 dengue hemorrhagic, 173 dengue fever, and 20 dengue shock syndrome). The proposed method achieved 0.847, 0.782, and 0.802 of accuracy, sensitivity, and specificity, respectively. This study is advantageous to predict the diagnosis of dengue disease using simple hematological and clinical parameters.

A heart disease prediction system has been proposed using a decision tree [72]. The structural information like age, chest

TABLE 3. Summary of Bayesian methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[66]	Vembanda-samy et al.	NB	2015	A dataset composed of 500 clinical records obtained from Chennai	Heart disease	Precision: 71%, Recall: 74%, F-measure: 71.2%	Heart Diseases Detection Using HDPS based on Naive Bayes
[63]	Wang et al.	Bayesian network, SVM, LR, and NB	2014	Dataset containing 50,000 cases of cancer patients in Taiwan	brain Metastasis Lung Cancer	Accuracy: 82%, Sensitivity: 83.28%, Specificity: 80.87%	predicting lung cancer using Bayesian network
[60]	Rana & Jyoti	NB	2014	Heart disease male dataset	Heart attack	—	Heart Disease Prediction using Naive Bayes Classification
[62]	Kononenko	inductive learning of DT, and NB	2014	Tumor(339), breast cancer(288), thyroid(884),and Rheumatology(355) from Ljubljana	Tumor, breast cancer, thyroid, and Rheumatology	Accuracy-Tumor: 49%, breast cancer: 78%, thyroid: 70%, Rheumatology: 67%	Inductive learning of decision tree and Bayesian for Medical Diagnosis
[61]	Aidaros et al.	NB, LR, DT, and NN	2012	15 datasets of various diseases from the UCI repository	Multiple diseases including cancer, hepatitis, and liver disorder	Accuracy: 97.43%, AUC: 99%	Medical data classification using NB, LR, DT, and NN.
[67]	Wei et al.	MANB	2011	1411 Alzheimer's patients	Alzheimer's	MANB Training Time: 16.1 s, AUC: 72%	MANB method to predict late-onset Alzheimer's disease from genome-wide data
[64]	Kazmierska et al.	Bayesian Network	2008	Brain tumour dataset of 142 patients	Brain Tumor	Accuracy: 84%, Specificity: 87%, Sensitivity: 80%	Application of the Naive Bayesian Classifier to optimize treatment decisions
[65]	Sakai et al.	Bayesian Network	2007	A dataset of 169 patients suspected of acute appendicitis	Appendicitis	—	diagnostic prediction of acute appendicitis using Bayesian network

TABLE 4. Summary of decision tree methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[78]	Sun & Zhang	Decision Tree	2019	5057 records of DR patients	Diabetic Retinopathy	Accuracy: 86.82	Diagnosis of DR based on EHR using five ML methods
[76]	Sowmien et al.	C4.5 Decision Tree	2016	155 patients (both healthier and live affected)	Hepatitis	Accuracy: 85.81%, Precision: 85.3%, Recall: 85.8%, F-measure: 83.9%, ROC: 76.9%	Diagnosis of hepatitis using C4.5 decision tree
[77]	Lungu et al.	Decision Tree	2016	72 patients with suspected PH	Pulmonary hypertension	Accuracy: 92%, Sensitivity: 97%, Specificity: 73%, PPV: 93.22%	Diagnosis of pulmonary hypertension from MRI images using decision tree
[74]	Tyagi & Preetvanti	Decision Tree & Fuzzy System	2014	30 Asthmatic patients	Asthma	Accuracy: 90%, Kappa: 78.32%	Diagnosis and control level of asthma using decision tree and fuzzy system
[72]	Sabarinathan & Sugumaran	J48 Decision Tree	2014	240 patients (120 each for healthy and heart-disease)	Heart disease	Accuracy: 85%	Heart disease prediction using decision tree J48 classifier.
[75]	Shouman et al.	Decision Tree with voting	2011	Cleveland Clinic Heart disease dataset with 76 attributes	Heart disease	Sensitivity: 77.9%, Specificity: 85.2%, Accuracy: 84.1%	Diagnosis of heart disease patients using decision tree with voting, equal frequency, discretization, and gain ratio
[71]	Tanner et al.	C4.5 Decision Tree	2008	1200 Dengue patients	Dengue fever	Accuracy: 84.7%, Sensitivity: 78.2%, Specificity: 80.2%	diagnosis of dengue fever by using C4.5 decision tree classifier
[73]	Kuo et al.	Decision Tree	2001	243 breast cancer images	Breast Tumor	Accuracy: 85.5%	Diagnosis of breast cancer using decision tree

pain, gender, and heart rate has been extracted using machine learning tools. The J48 decision tree classifier has been used for feature selection evaluated on a dataset of 240 instances (120 each for healthy people and heart disease affected people). The proposed method achieved 0.85 of accuracy accurately predicting the conditions of heart disease.

In [73], Kuo *et al.* have been used decision trees for diagnosis of breast tumour from ultrasonic images. The training of the proposed model has been performed was extracted from the region of the interest (ROI). A 24-D texture feature set has been used to classify the tumour as malignant or benign. A dataset composed of 243 breast tumour images was

collected from 1997 to 1998. The empirical results of this study have shown 95.5% accuracy.

In [74], the decision tree has been suggested for asthma diagnosis, and a fuzzy system was used to measure asthma control level. Asthma has been diagnosed using symptoms such as a sore throat, dry cough, and sneezing etc. while the control level was estimated based on breath shortness, daytime symptoms, and action limits etc. The data was collected in the form of questionnaires from the patients. Decision tree classifier was used for the diagnosis of asthmatic patients and achieved 0.90, and 0.783 of accuracy and kappa coefficient, respectively.

Shouman *et al.* [75] have been proposed decision tree algorithm with additional parameters like voting, equal frequency, discretization, and gain ratio for diagnosing heart disease patients. The proposed method was evaluated on Cleveland Clinic Heart disease dataset (composed of 76 raw attributes in which 13 are selected for this study). The result of this method was compared with commonly used decision tree classifiers such J4.8 and Bagging algorithm. The proposed method achieved 0.779, 0.852, and 0.841 of sensitivity, specificity, and accuracy, respectively.

The author has been suggested C4.5 decision tree classifier for the diagnosis of Hepatitis [76]. A total of 155 patients (both healthier and the liver affected patients) were selected in this study. There were 19 attributes used to construct the method including age, sex, steroid, fatigue, and varices etc. The proposed method achieved 0.8581 of accuracy for hepatitis diagnosis.

Pulmonary hypertension (PH) has been diagnosed from right heart catechization (RHC) and magnetic resonance imaging (MRI) using decision tree algorithm [77]. A total of 72 patients with suspected PH underwent MRI and RHC, in which 57 has PH, and 15 were diagnosed with no PH. The proposed model correctly classified 92% of the PH patients, while the rest of them were considered as misclassification error. The optimal results of this study can decrease the need for RHC with suspected PH.

In [78], five machine learning methods are employed to diagnose diabetic retinopathy (DR) from EHR data. A large retinal dataset comprised of 5057 records is collected from 301 hospitals in China. The preprocessing steps such as label binarization, standard sceleration, and normalization of values are performed to improve DR disease diagnosis accuracy. SVM, NB, DT, LR, and RF classification models are used for classification tasks. DT model achieved high diagnostic accuracy of 86.82%. The proposed method is advantageous as compared to other DR diagnostic methods because of low cost, low threshold, and high diagnostic accuracy.

C. DEEP LEARNING METHODS

Deep learning (DL) also referred to as deep machine learning or hierarchical learning, is used to model a high level of abstraction in data by applying various processing layers. Over the last few years, diagnosis of disease is performed

using traditional methods including rule-based methods, statistical learning methods, and machine learning methods such as SVM, NB, and RF [79]. With the advancement in DL approaches, it is widely adopted by the researchers over the globe for various disease identification. DL has shown significant performance in many domains by capturing extensive range dependencies and constructing dense hierarchical features in an efficient way [80]–[83]. In order to build a construct effective disease diagnosis system, advancement in both data representation and development of ML architectures are imperative [84], [85]. Deep learning is based on deep neural networks (usually the term deep is used for networks for more than one hidden layer) which includes various networks such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Auto-encoder (AE), Deep Belief Networks (DBN), and more. We study the various deep learning techniques employed on EHR data for clinical tasks and discuss their practical advantages and future directions.

1) CONVOLUTIONAL NEURAL NETWORKS (CNN)

CNN has become very popular in medical diagnosis over the last few years. It is widely used in medical imaging data including lab reports and pathological reports etc. It imposes local connectivity on the unstructured data. The convolution layer and pooling layers are used in the experimentation of CNN architecture. CNN is used to extract highly informative features from EHR [95], [96].

A CNN based method has been used for the prognosis of diabetic retinopathy [87]. This CNN model has been restructured to build a new BNCNN model by adding the BN layer to traditional LeNet model. The dataset containing 500 records of DR, and no DR is gathered from 301 hospitals in China and abroad during the time period of 2009 to 2013. The proposed has been achieved state of the art training accuracy of 99.85% and testing accuracy of 97.56% which is 2% higher than LR method. BNCNN model is effective for preventing the gradient diffusion, improving the training speed and accuracy of the model. A model-based reasoning (MBR) algorithm for disease diagnosis based on EMR and natural language processing showed 95.86% accuracy [97].

Mehrabi *et al.* [93] have employed deep learning techniques for temporal pattern discovery. The dataset of Rochester Epidemiology was used for this research from Olmsted country. They have modelled every patient's data as a temporal matrix in which rows contains diagnosis coded of ICD-9 and Healthcare Cost and Utilization Project and clinical classification software (HCUP-CCS), and column contains diagnosis years with age not more than 18 had been selected. They constructed a Boltzmann machine network with three hidden layers and used the temporal matrix as a visible node. The results of the proposed method have shown relationships, i.e., blindness and eye disorder diagnosis code. The network weights were explored as a feature in the patient's data. The additional patterns marked by medical experts can be used to gain better results.

TABLE 5. Summary of CNN methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[86]	Mullenbach et al.	CNN and Framingham risk score	2019	4 datasets of 4298 patients collected from grade-A hospital in China	cerebral infraction(CI), pulmonary infarction(PI) and coronary heart(CH)	Accuracy-CI:96.5%, PI:95.6%, CH:93.6%	Multiple Disease Risk Assessment With Uniform Model Based on Medical Clinical Notes
[87]	Y Sun.	BNCNN	2019	500 samples of DR patients	Diabetic Retinopathy	Accuracy: 97.56%	Diagnosis of Diabetic Retinopathy using neural network of one-dimensional convolution
[88]	Shi et al.	CAML, DR-CAML, and CNN	2019	MIMIC-III UCI repository containing 47724 discharge summaries of 36,998 patients	Multiple diseases	Macro AUC: 89.7%, Micro AUC: 98.5%	Convolutional Attention for Multi-Label classification based on CNN used to predict medical codes from EHR
[89]	Miotto et al.	unsupervised deep feature learning	2016	700,000 patients from the Mount Sinai data warehouse	78 diseases	AUC: 77.3%, Accuracy: 92.9%, F-score: 18.1%	Deep Patient: to derive patient representation from EHR that facilitates predictive modeling
[90]	Choi et al.	NN with GRU	2016	3884 HF cases	Heart failure	AUC: 88.3%	Prediction of heart failure onset CNN with gated recurrent units
[91]	Cheng et al.	Basic, Late, and Early Fusion CNN	2016	EHR containing 319,650 patients data	CHF, COPD	AUC-CHF: 76.75%, COPD: 73.88%	Deep learning framework for analyzing patient EHRs
[92]	Zhu et al.	RV, dConv, and CNN	2016	A dataset of 218,680 patients	COPD, Diabetes, HF, and Obesity	Rand Index: 98.87%, Purity: 98.82%, normalized mutual information: 95.16%	Measuring patients similarities from EHR using CNN
[93]	Mehrabi et al.	Deep Learning	2015	Rochester Epidemiology Project with 502,820 patients	Multiple Diseases	—	Temporal pattern discovery over REP project of ICD-9 diagnosis code
[94]	Nie et al.	Sparsely Connected Deep Learning	2015	900 disease concepts	Multiple diseases like Parkinson, Asthma, Lyme, and Liver	Accuracy: 98.21%	Disease inference from health-related question using deep learning

A novel approach of disease inference has been performed by asking health questions, selected possible disease of their symptoms, and then applied sparsely connected deep learning method to infer possible diseases by giving the questions of patients [94]. Two main features were used in this research. Firstly, they mined discriminant features from raw medical features. Secondly, they considered raw features and signature in the first layer and hidden node in the second layer, respectively. The proposed method is sparsely connected, and node numbers are adjusted automatically, unlike the traditional deep learning methods, which are densely connected, and node numbers are tediously adjusted. This method is not suitable to identify discriminant features for each disease, which can be observed in future studies to gain wide accessibility of this method.

Deep patient: an unsupervised deep feature learning method introduced in [89] by Miotto *et al.* They employed this method for predictive modelling using clinical notes by driving patient representation. A Mount Sinai dataset of 7,000 patients was used by capturing hierarchal regularities and dependencies in clinical notes. The evaluation was performed

on 76214 patients comprising 78 diseases. The results of this study have been outperformed those methods which were achieved using a representation based on raw health records. They achieved top performance to predict various diseases including cancer, diabetes, and schizophrenia. DL methods can improve clinical predictions if derived patient representations are used with this. The laboratory results can be included in this study to enhance its representation performance.

Choi *et al.* have proposed a novel method for signifying diverse medicinal concepts as real-valued vectors and deep learning method has been used to construct patient representation [90]. The trained medical concepts were utilized to capture heart failure prediction. The training of various models has been performed on 3884 and 28903 controls of HF. The classification methods (LR, SVM, NN, and KNN) achieved 23% AUC improvement using proposed representation. The effective results can be acquired by adopting lab reports and patient demographic information along with this representation.

A CNN model has been applied to analyze patient information from EHRs [91]. EHR of every patient was represented

as a temporal matrix and CNN model was formed by creating four layers such as input layer (contains EHR matrices), convolution layer (extract phenotype from input layer), max pooling layer (presenting sparsity on identified phenotypes), and fully connected layer (softmax prediction). Moreover, temporal fusion procedure was examined for the smoothness of patient EHR. The proposed was evaluated qualitatively and quantitatively on EHR dataset collected over 4 years. CNN model can be avoided from over-fitting by reducing the number of parameters in the future. The proposed method can also be used for other works.

Zhu *et al.* [92] measured patient similarities using the DL model along with medical concepts. The similarities were evaluated based on the temporal matching of longitudinal patient clinical notes. Both the supervised and unsupervised methods were employed for this task, which preserves temporal properties in EHR. CNN model was used as a supervised learning method and learned representations of clinical records with medical events embedding from word2vec. The results of this study have been outperformed baseline methods with values of R1, Purity, and NM1 are 0.9887, 0.9882, and 0.9516 respectively. The future work includes the addition of time interval information to resolve data regulatory issues and the implementation of the proposed method in other domains such as health visualization.

Convolutional Attention for Multi-Label classification (CAML) based on CNN was employed to predict medical codes from electronic health records [88]. CAML used the attention mechanism to pool the convolution output for every label. The proposed method has shown strong improvements for the prediction task over the baseline methods of ICD9code. CAML was designed as an adaptable method for ICD-10 code. The discharge summaries integrated with MIMIC-III dataset (open repository for ICU medical records), can be used in the future to handle non-standard writing and out of vocabulary tokens.

The four multi-label classification approaches were employed by assigning ICD codes to discharge summaries, MIMIC-II and MIMIC-III clinical datasets [98]. They evaluated their models on all the ICD-9 codes, to make sure the real-world ICD-9 tagging. The experimentation was performed on the SVM model, continuous bag-of-words (CBOW) model, CNN model, and HA-GRU ('bidirectional Gated Recurrent Unit model with a Hierarchical Attention Mechanism'). HA-GRU achieved promising results by applying tokenization and hierarchical segmentation.

Brown *et al.* [99] have been used deep convolutional neural networks to diagnose retinopathy of prematurity (ROP) from retinal images. A dataset composed of 5511 retinal images and 5-folded cross-validation was used to evaluate the method. A standard reference diagnosis (RSD) has been assigned to every image based on image-grading (3 experts) and clinical diagnosis (1 expert) and classified into the normal, pre-plus, plus disease. This method achieved 0.94 and 0.99 of AUC for normal and plus disease, respectively. Furthermore, the algorithm achieved 0.91 of accuracy on a test

set of 100 images, which has outperformed 6 of 8 ROP experts.

The disease prediction model based on electronic medical records has been developed to assess multiple diseases [86]. The proposed method has been characterized by the convolutional neural network for multiple diseases prediction. This method has been evaluated using data of 4298 patients with cerebral infection, pulmonary infection, and coronary heart disease. The CNN algorithm achieved the highest of 96.5% accuracy, and 96.6% F1-measure on D1 dataset for cerebral infection disease.

2) RECURRENT NEURAL NETWORKS (RNN)

Recurrent networks belong to the category of artificial neural networks. Unlike feedforward neural networks, the connections in nodes create a directed graph which displays a time sequence for temporal dynamic behaviour. The accurately mining of EHR records is very challenging because of various factors including incomplete prescription, free text writing, and disease variability.

A RNN based long short term memory (LSTM) architecture has been proposed for predicting the risk of heart failure from EHR [100]. A real-world combined dataset of 20,000 patients of congestive heart disease was to evaluate the performance of LSTM method. LSTM has outperformed LR, RF, and AdaBoost algorithms after 5-fold cross validation.

Deep-Diagnosis [101] was introduced by Shi *et al.* to address these challenges. The proposed method was constructed using RNN-based prediction diagnosis algorithm by mining pediatric EHR data. They performed preprocessing of Chinese unstructured EHRs dataset feature vector has been created using NLP concepts by transferring these values into sentence vector. The patients' symptoms and their relationship have been identified using the proposed bidirectional Recurrent Neural Network. A pediatric dataset composed of 81476 records were used to train and test the model which achieved 80.912 of precision. The method can be cross evaluated on various EHR datasets along with some other measures of NLP can be used to extract specific information from EHRs data as a future work of this study.

Luo [105] has been used RNN based Long-Short Term Memory (LSTM) model for classifying relations in EHRs data. The model was evaluated on relation classification challenge database i2b2/VA. The proposed model achieved f-measure of 0.61, 0.80, and 0.68 for classifying medical problem-treatment, medical problem-test, and medical problem-medical problem relations. The comparison between segment and sentence LSTM has been performed to demonstrate the difference between context and concept text. They evaluated the word-embedding impact on the performance LSTM and showed that the performance of the proposed method is comparable with existing methods.

The diagnosis and probability of clinical visits have been predicted using demographic information and RNN method (RNN-INFO) [102]. The data dimensionality has been reduced using neural network hidden layer and by extracting

TABLE 6. Summary of RNN methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[100]	Jin et al.	LSTM-RNN	2018	Real-world dataset of 20,000 patients	heart disease	ROC: 64.83%, F1 score: 27.87%	Predicting the risk of heart failure from EHR using LSTM-RNN model.
[101]	Shi et al.	BiRNN	2018	Xiamen dataset containing 81,476 pediatric EHRs	8 pediatric diseases	Precision: 80.91%	Deep diagnosis: patient's symptoms and their relationship identified using BiRNN.
[102]	Wang et al.	RNN-INFO	2018	289,000 clinical notes of 28,000 patients	Multiple diseases	F-score: 85.39%	detection of diseases Using RNN and Demographic Information
[103]	Wu et al.	RNN	2018	4013 patients of Olmsted County Birth Cohort, 4000 patients of Physionet	Pediatric Asthma	Precision: 84.54%, Recall: 85.65%, F-measure: 85.08%	Classification of pediatric asthma (chronic disease) using RNN.
[104]	Chu et al.	CA-RNN	2018	8845 patients of cardiovascular diseases	cardiovascular	Recall: 92.9%, Precision: 56.1%, F-measure: 82.1%, AUC: 96.2%	Using Context-Attention based RNN to detect adverse medical events from EHR
[105]	Luo	LSTM-RNN	2017	data from the 2010 i2b2/VA challenge	Chronic kidney disease	F-measure: 80%	RNN for classifying relations in clinical notes
[106]	Dong et al.	Bi-directional LSTM RNN	2017	992 Chinese EMRs (500 discharge summaries and 492 progress notes)	Multiple diseases	Precision: 91.23%, Recall: 92.45%, F-Measure: 91.83%	Transfer Bi-directional LSTM RNN for Named Entity Recognition in Chinese EMR
[107]	Esteban et al.	RNN	2016	20161 patients of Charite' Hospital in Berlin	Kidney	AUC: 83.3%	Predicting Clinical Events by Combining Static and Dynamic Information using RNN

the hidden representation in clinical notes. A dataset composed of 289 thousand medical records and 28 thousand patients were used to evaluate the proposed model. This study has shown promising results than non-temporal models on both diagnosis and probability prediction. Furthermore, they verified that RNN-INFO (Model with Demographic Information) had shown better results than only RNN (Model without Demographic Information). The model with demographic information has a significant performance than a model without demographic information in probability prediction.

Wu *et al.* [103] introduced event sequences and their properties (evenness, synchronicity, and co-cardinality) for the classification of pediatric asthma (chronic disease). They also determined how the inverse of these properties, i.e. uneven, asynchronous, and multi-cardinality can assist accountings of relative time. The results of this study have shown that embedding time stamp with RNN model gives a better classification of patients with no asthma rather than asthma patients.

A bidirectional recurrent neural network (BiRNN) has been proposed to extract automatic medical information such as diseases and treatments from Chinese EHRs data [106]. The proposed model has been constructed using two steps. First steps involved the training of shallow BiRNN model. In the second step, medical information has been transferred from the general domain to train BiRNN for automatic recognition of concepts. Transfer-BiRNN has shown promising

results than baseline methods and achieved 88.7%, 88.49% accuracy of disease and symptoms prediction.

A method based on RNN and the combination of static and dynamic information has been proposed for the prediction of clinical events [107]. A database collected from Charite Hospital Berlin of kidney transplanted patients was used to evaluate the method. The database has three endpoints, i.e. kidney rejection, patient death, and kidney loss. The proposed model was used to predict these endpoints in each patient after 6-12 months of the clinic visit. RNN with Gated Recurrent Units (RNN-GRU) has outperformed other methods for this task. Furthermore, optimum results have been obtained using binary encoding inputs rather than normalizing input data. The other kidney loss information like biopsies and its results can be used in this model to obtain better results in the future.

The challenge of adverse medical detection (ADEs) has been solved using the proposed recurrent neural network based on context attention embedding (CA-RNN) [104]. The salient words concerning target have been located using the proposed context-aware attention technique. Then they combined a deep learning approach with the proposed technique to boost the performance of adverse medical event detection. The proposed method was evaluated on 8845 cardiovascular medical records and achieved 0.929 and 0.561 of recall and precision for AME of Ischemia. The future work of this study may include larger dataset and evaluation of this method on other diseases.

TABLE 7. Summary of DBN methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[108]	Hao et al.	MD-RCNN, DBN	2019	31,919 patients data of Wuhan hospital, China	Multiple diseases i.e. Coronary heart, and Pulmonary infection	Accuracy: 96%, Recall: 98.08%	disease risk prediction using multimodal data-based RNN and DBN
[109]	Farzi et al.	DBN with Greedy Approach	2017	New York University (263 samples) and Neuroimaging (73 samples) dataset	ADHD	Accuracy- NYU: 63.68%, NI: 69.83%	Diagnosis of Attention Deficit Hyperactivity Disorder (ADHD) Using DBN on Greedy Approach
[110]	Faturrahman et al.	DBN, SVM	2017	98 images [49 of AD patients, and 49 of healthy people]	Alzheimer's Disease	Accuracy: 91.76%, Sensitivity: 90.50%, Specificity: 92.96%	Classification for Alzheimer's Disease detection using DBN
[111]	Al-Fatlawi et al.	DBN	2016	31 people Parkinson's Data Set (23 with Parkinson's disease and 8 healthy) collected from the University of Oxford	Parkinson's Disease	Accuracy: 94%	Efficient Diagnosis System for Parkinson's Disease Using DBN

3) DEEP BELIEF NETWORK (DBN)

A deep belief network is one of the representations of deep learning [112]. It has been widely adopted for multiple tasks such as pattern recognition, handwriting recognition, speech recognition and many other tasks [113], [114]. It is also used for classification task [115], soft sensor [116], and image-based monitoring [117]. DBN is an unsupervised method which works on unlabeled data. It is mostly applied to construct hierarchical structures and uses unsupervised learning for feature representation. It is restricted in acquiring features because there is no intermediate relationship between the neurons of the same layer [118].

In [108], a new multimodal disease recurrent convolutional neural network (MD-RCNN) has been proposed for the risk prediction of disease. This model has been used to extract both the structured and unstructured features in fine-grained. Deep Belief Network (DBN) has been used to obtain a non-linear connection between unstructured and structured data by fusing the features. The model has been evaluated on two Chinese datasets from 2013 to 2015 and achieved a state of art accuracy of 96%. The future directions of this article include other diseases prediction and people behaviours prediction.

A deep belief network (DBN) has been used in [111] for the diagnosis of Parkinson's disease (PD) based on the speech signals obtained from the UCI repository. The proposed method was trained on various healthy and patient voices, and feature extraction was performed by inputting DBN. The proposed method was used to categorize PD which composed of one output layer, and two stacked restricted Boltzmann machines. The RBM model was employed as unsupervised learning to overcome the problems of initial weights and backpropagation used supervised learning for fine-tuning. The proposed method achieved 94% of accuracy for the diagnosis of Parkinson's disease.

In [109] DBN has been employed for the diagnosis of one of the most common disease namely attention deficit hyperactivity disorder (ADHD). The proposed method used a greedy approach for the construction and the training of the network. The two training and testing datasets (New York and Neuroimaging) were provided by Global competitions ADHD-200. There were (222, and 41) and (48, and 25) samples of training and testing samples of New York University (NYU) and Neuroimaging (NI) dataset, respectively. They achieved state of the art accuracy of 0.6368 and 0.6983 on NYU and NI datasets respectively.

Faturrahman *et al.* [110] have been proposed DBN method for the classification of Alzheimer's disease (AD) based on structural modalities (MRI data). The result of this model has been compared with the support vector machine (SVM). The feature obtained from mean and standard deviation calculation (MSD) with fine-tuned hyper-parameters such as hidden nodes 25, epoch 250 and momentum from 0.5-0.9 of the proposed method achieved 73.6%, 71.2% and 76% accuracy, sensitivity and specificity, respectively and the based on voxel value (VV) with epoch 100, and momentum 0 achieved 91.7%, 90.5%, and 92.9% accuracy, sensitivity, and specificity, respectively. AD detection can be improved by applying feature selection and modality techniques like PET and CSF.

4) AUTOENCODER (AE)

Autoencoder is the representation of the artificial neural network, which is based on unsupervised learning to learn data encodings. It is generally used for dimensionality reduction by learning to encode for a set of data. AE concepts are now employed to learn a generative model of data [119], [120]. The modern concepts of AI have included AE stacked in deep neural networks [121].

Stacked autoencoder and softmax classification methods have been used for the diagnosis and classification of cervical

cancer [122], [123]. A UCI dataset composed of 668 samples, 4 targets and 30 attributes were used for training and testing of the method. The proposed methodology was divided into two steps: autoencoder was applied on raw data to obtain reduced dimensionality dataset in the first step, and softmax layer has applied for the classification on the trained dataset in the second step. The dataset was divided into a training set (70%), and a test set (30%). The proposed model was implemented in four target variables (Cytology, Schiller, Hinselmann, and Biopsy) and their classification performance was compared which achieved 0.978 of correct classification rate. The training of this model takes too much time because of the dimensionality reduction of samples. Therefore, future measures can be taken to improve the training time of the model.

In [124] automatic encoding of medical procedure has been done using autoencoders. The classification of medical procedures has been performed by query matching characterized by convolutional neural networks. The pipeline was created using CNN and AE with logistic regression (LR). F1 Micro score of 0.702 and 0.608 achieved on autoencoders and convolutional neural networks, respectively to determine relevance between query text and category text. A comparison of the proposed method with various existing methods has been performed based on the suitability of automatic encoding. The Bayes algorithm can be adopted as a guiding procedure for knowledge reasoning. The current dataset (only 24,092 pairs) is small for fine-tuning, which can be expanded as future work.

Hwang *et al.* [125] have been performed performance comparison of disease prediction by using generative adversarial networks (GANs) and conventional networks in combination with missing value prediction techniques. The promising results have been achieved using stacked autoencoder (missing value predicting method) and with auxiliary classifier GANs (AC-GANs: disease prediction) with 0.98, 0.95, and 0.99 of accuracy, sensitivity, and specificity, respectively. GAN is used as a generic model only, and AE is used to fill missing values in this study. The future direction of this work includes using GAN to fill the missing values.

The automatic diagnosis of prostate cancer (PCa) from MRI images has been performed using random forest and autoencoder methods [126]. An imbalance dataset composed of diffusion-weighted, B-value, apparent diffusion coefficient, and transaxial T2 weighted images were collected from 2016 challenge of PROSTATEx. Single-level Sparse Autoencoder (SAE) has been applied to extract high-level features. The class imbalance problem has been resolved using Weka resampling, SMOTE ('Synthetic Minority Over-sampling Technique') and ADASYN ('Adaptive Synthetic'). The proposed method has been compared with other baseline methods, and it has observed that ADASYN followed by random forest classifier achieved promising results with 0.979 area under the ROC curve, 0.936 of accuracy, and 0.94 of F-measure for the diagnosis of PCa. A fully automated diagnosis system of prostate cancer can be built in future.

DL FRAMEWORKS FOR DISEASE DIAGNOSIS SYSTEMS

Internet is producing bulk of raw data which can be effectively used to train DL models. Because of the limitations of traditional ANN and ML techniques, DL leverages to cope with larger dataset, make intelligent decision and disease diagnosis, etc. The recent advancement in key enablers influence DL based disease diagnosis such as efficient regularizers (L1/L2, Dropout, etc.), loss functions (ADAM, SGD, RMSProp, and so on), and commodity hardware (e.g. GPUs and I/O bus speed). Moreover, GPU based systems have much improved to execute DL models for various diagnosis of diseases. The technology giants like Google, Microsoft, Apple and Facebook are massively investing on GPU-accelerated DL platforms. The DL models in these libraries are used to carry out the training with the help of data and model parallelism. The details of prominent platforms used for disease diagnosis are as follows:

- 1) TensorFlow: It is a free and open source library developed by Google engineers. It provides an interface to access recent ML and DL algorithms. It can be accessible by individuals (CPU-based) or through large-scale distributed systems (GPU-based). It is a symbolic math library written in Python and C++ languages. Tensorflow API is widely used for various disease diagnosis such as cataract, glaucoma and cancer detection.
- 2) Torch: It is an open source ML library written in Lua programming language. It provides vast range of DL algorithms with fast and efficient GPU support. It is widely adopted for social media applications such as Facebook, Twitter and Google, etc. Unlike Tensorflow, the use of Torch library for disease diagnosis is rather limited.
- 3) Caffe: It is the acronym of convolutional architecture for fast feature embedding, and was developed by Berkeley Vision and Learning Center. It provides modular and robust framework for deep learning algorithms by exploiting multiple GPUs. Caffe uses JSON-like text to describe network and offers various prototypes in speech, vision and multimedia. It is anticipated to utilized for large-scale application including disease diagnosis systems.
- 4) Deeplearning4j (DL4J): IT is an open source with a GPU support library. It was developed by Karapathy and written in Java and Scala programming languages. This framework combines well with Spark [127] and Apache Hadoop [128] by employing multiple CPUs/GPUs. It is not enormously used in medical imaging and appears to be declining in the medical field. Its speed is comparable with Caffe framework by exploiting multiple GPUs. DL4J exhibits better performance as compared with Torch and Tensorflow.
- 5) Theano: This library is written in Python language, which leverages numpy code and improves performance as compared to basic python. It is based on symbolic logic which is challenging to learn and create

TABLE 8. Summary of autoencoder methods.

Ref.	Author(s)	Methods	Year	Dataset	Focused Disease(s)	Performance Measures	Objective
[122]	Adem et al.	Stacked AE and Softmax classification	2019	UCI dataset with 668 samples	Cervical cancer	Accuracy: 97.25%	Classification and diagnosis of cervical cancer with stacked autoencoder and softmax classification
[124]	Deng et al.	CNN and AE	2018	ID DIACOS Berlin dataset with 24,092 query-category training pairs	Multiple Diseases	Precision: 85.21%, Recall: 59.56%, F1-score: 70.29%, Specificity: 95.3%	Automatic encoding of medical procedures using CNN and autoencoders
[126]	Abraham & Nair	Sparse AE, and RF	2018	TCIA Cancer Imaging dataset composed of 330 lesions	Prostate Cancer	ROC: 97.9%, Kappa: 87.3%, Accuracy: 93.65%, F-measure: 94%	Diagnosis of prostate cancer from MRI images using sparse autoencoder and random forest classifier
[125]	Hwang et al.	Stacked AE and GAN	2017	569 records of breast cancer with 357 for benign and 212 for malignant cancer	Breast cancer	Accuracy: 98.05%, Sensitivity: 95.28%, Specificity: 99.47%	Disease Prediction from EHR Using stacked autoencoders and Generative Adversarial Networks (GAN)

complete solutions. It is generally used for evaluating and manipulating mathematical expressions such as matrix values and can run effectively on CPU and GPU architectures.

- 6) WILL: It is a deep learning framework based on C++ language. It is also compatible with other interface of Python, C, and assembly language. WILL has cross-platform compatible with all operating systems. It is more convenient in configuration, and very rarely used for disease diagnosis. The performance and speed of WILL is comparable with DL4J and Caffe framework.
- 7) Keras: It is an open source library written in Python language. It runs on the top of Tensorflow or Theano. It is most widely adopted by the researchers for deep learning associated tasks. It is very simpler to use and to build complete solutions. Though, it uses both the Theano and Tensorflow as backend, a careful design is required to select the dimensions of input data. It has shown significant performance various DL tasks including computer vision, speech recognition, and multimedia, etc. Over the last few years, Keras is used for the detection and classification of various diseases.
- 8) CUDA: It is a parallel computing platform and application program interface developed by Nvidia. It permits software developers to use CUDA enabled-general purpose GPUs (GPGPU). The built-in deep neural network library in CUDA is used to performed to standard deep learning routines such as convolution, batch-normalization, pooling, and activation of layers, etc. It is one of the most widely used framework for various image processing tasks.

It can be concluded from the above discussion that the researchers have access to wide range of deep learning libraries. These libraries can be implemented to solve various research domain problems. However, deep learning applications in disease diagnosis have not been surveyed

extensively. The open problems related to these technologies are described in the subsequent section.

D. OPEN PROBLEMS

Although electronic health records assist both the patient and clinicians to keep track of patient progress, and provides remote and faster access, still there are lot of challenges associated with EHR such as data volume, staff training, data alteration, vendor selection, insufficient support material, implementation delays, scarce info, data security and other Medicare technical issues. Different methods have been proposed to overcome the challenges associated with EHR. Though these methods are not intelligent enough to tackle all the problems of EHR, several successes have been reported by employing machine learning and deep learning methods.

Disease diagnosis using rule-based methods is the most commonly practised techniques by medical specialists. The rule-based methods are grounded in production rules (condition-action pairs). The construction of complex rules in case of highly dependent data and excessive volume is a very laborious and time-taking effort. Another challenge associated with rule-based methods is the scalability of data, because once you have written the rule, the addition of anything beyond that rules may lead to possible errors. It is always ambiguous to choose the rule when two different rules are created for the same problem.

Diagnosis using machine learning methods can work when condition reduced to classification task on physiological data, in areas where the clinician can be able to identify patterns indicating the presence or absence of that condition. ML methods don't work properly if it is not a classification problem. Usually, it requires a large amount of data for the training of the model; we cannot train the model with a small amount of data. Data preprocessing is another main challenge, because to deal with missing values and to improve the image quality is very strenuous (to obtain the minute details, and to differentiate between the background and the

object). The designing of the ML model with labelled data is often useful, but the development of unsupervised ML model is very tedious and lengthy. Other challenges include data accessibility, accurate result interpretation, pertinent usage of image structure to design a model, and application of results in clinical practice. The results of disease diagnosis using ML methods can be improved by dealing with these challenges.

Deep learning methods have achieved high performance for medical image analysis task. These methods are easily scalable, independent of deep knowledge and can be developed rapidly. One of the major obstacles with a deep learning approach is the lack of training dataset. The acquisition of pertinent annotation and labelling is also the main problem. Labelling of the larger dataset (often annotated by the domain experts) such as vessel segmentation and 3-D radiology segmentation is very problematic because it takes a significant amount of time for labelling. The labelling noise and class imbalance are considered as the open challenges for researchers.

III. DISCUSSION

The physicians or medical experts diagnose a specific disease by observing the patient's physical condition, exploring symptoms, or through laboratory tests and pathological reports. The patient's history is usually stored as a prescription, for necessary medication, treatment plans, and to track disease progression. Initially, it was done through paper form, which was not an efficient method because of many reasons including remote access, a reminder of vaccination dates, streamline workflow, and keep track patient's performance. With the advancement of technology, the prescriptions are saved in an electronic format (digital format) known as EHR or EMR. These electronic records give assistance to the physicians or medical experts to access the patient records instantly, to keep track of patients' due dates for checkups and immunizations and monitor patient's health performance.

Initially, some medical expert systems are designed for automatic diagnosis of diseases. They were generally based on queries, and the list of diseases was stored in the database against their symptoms. Medical expert systems are generally concerned on the computer writing programs to perform disease diagnosis. Although many applications have been developed, these systems are still very expensive, and there have been very technical and usability issues. There are various issues related to the acceptance of these medical systems as discussed in [88], [92], [129]. These issues are generally concerned with accepting a ratio of these systems, more time consuming, and manufacturing cost etc. Fuzzy expert systems are commonly used for both experimental and applied medicine. Pabbi [130] employed the expert system to predict dengue fever and achieved 95% accuracy. They used a sample based dataset along with lab reports and symptoms to evaluate the proposed method. Main and Oldham [131] identified 17 rules for designing expert systems. They further highlighted the importance of these rules and changing the behaviour of the people with time explored. Knowledge-Based (KB) systems

are widely adopted systems to handle knowledge-intensive tasks. They are used to handle the problems related to laboratory and pathology. A survey of KB systems in the laboratory was carried out by Spackman and Connelly [132], suggested that these systems have increasing values in laboratory and pathology.

An EHR contains various information including X-ray, CT, pathology and retinal images, laboratory results, progress notes, and medical histories. This information has been used extensively for automatic detection and classification of diseases. Although, this information is not complete enough to accurately diagnose the disease, different measures are used to achieve the optimum diagnostic accuracy like preprocessing, feature extraction, and implementation of the model. The textual information from patient progress notes and medical histories have been extracted for disease diagnosis [25], [133], [134]. The retinal images are used for various retinal diseases such as glaucoma [135], diabetic retinopathy [136], [137], hemorrhage [138], hard-exudate [139], retinopathy of prematurity [140], and age-related macular degeneration detection [141] etc. The pathology images are used as an input for the detection of colorectal cancer [142], prostate cancer [143], breast cancer [144], [145], malaria [146], and metastases [147] and classification of various diseases including colon cancer [148], glioma [149], cancer tissue [150] and thyroid cytopathology segmentation [151] etc. Likewise, the multiple diseases have been diagnosed over the years based on x-ray and CT images i.e. thoracic disease [152], dental disease [153], [154], Alzheimer's [155], pulmonary embolism [156], lung disease [157], and the liver image classification [158].

Rule-based systems are used to extract medical information from electronic health records [25], [26]. These rules are based on conditional statements to match the relevant cases. These rule-based systems extract the attributes related to the disease. The human specified extraction rules are based on pattern matching scheme. Many rule-based systems have been developed to process clinical reports including MedLEE [36], MetaMao [159], and SymTex [160] etc. MedEx: a rule bases system was developed by Xu *et al.* to extract medical information and achieved 0.90 f-measure on extracting drug names, routes, dosages and drug frequencies from clinical notes [161]. Although many rule bases system has shown state of the art results with above 80% accuracy to diagnose a particular disease but to write these rules, is a very tedious job. Moreover, these systems are not intelligent enough to make a decision, if some symptoms are coherent or beyond the rules list.

With the advancement in AI concepts, Machine Learning method has been evolved very rapidly, and have been used for medical diagnosis over the last few years. ML technique is used to develop automatic, sophisticated, and objective methods for the analysis of multidimensional biomedical data [162]. The various supervised and unsupervised machine learning methods including SVM, NB, DT, and Random Forests are used in medical diagnosis [47]–[52], [54]–[57].

A review of Malaria diagnosis using ML methods has been performed by Sajana and Narasingarao [163]. Malaria disease has been identified using segmentation of the erythrocytes, life stages and density estimation of parasites, and by applying multiple techniques. These methods have shown promising results even on severe diseases like various types of cancer, diabetic retinopathy, and diabetes etc. There is still plenty of research gap is available for the researcher to improve the prediction results by introducing novel techniques or by integrating multidimensional heterogeneous data with multiple approaches of feature selection and classification.

DL methods are used to model a high level of abstraction in data by applying various layers. Diagnosis through deep learning works when it is limited to the classification task on EHR data. Various deep learning models such as CNN, RNN, RBM, and AE have been used medical diagnosis from EHR data [88], [91], [92], [98], [99]. These methods have attained worldwide acceptance to be applied in disease diagnosis from discharge summaries because of various layers including pooling layer, convolution layer, normalization layer, and fully connected layers. DL is considered as the best choice for medical image analysis including object detection [164], [165], image classification [166]–[168], registration [169]–[171], and segmentation [172]–[175] etc.

In this review, we observed the various approaches used for disease diagnosis from electronic medical records. We categorized these approaches into rule-based methods, machine learning methods, and deep learning methods. We have covered the major categories of machine learning methods including support vector machine, Naive bays, decision tree, and random forest methods. Deep learning methods are sub-divided into convolutional neural network methods, recurrent neural network methods, autoencoder methods, concept representation methods, and restricted Boltzmann methods. We covered various aspects of research in a tabular format, and also present a graphical representation of publications with respect to year and disease.

We tried our level best to cover all the latest and existing techniques used for disease diagnosis using EHR. The advantages and drawbacks of various proposed algorithms in each category are presented in this paper. The yearly decomposition of papers with respect to its category is also given in graphical format. We expected to give the professional a structure for the current research and to familiarize individuals with the various disease diagnosis approaches.

IV. CONCLUSION

A lot of work has been done for automatic extraction of useful information from electronic health records, clinical notes and discharge summaries. The physician uses features extracted information as an input for the automatic diagnosis of disease. Knowledge bases initially performed this extraction. In the present era, various rule-based learning, machine learning, and deep learning concepts are used for the extraction and diagnosis of diseases. There are several challenges

associated with this automatic extraction including missing values, incomplete information, and data abundance. We have reviewed recent research for the automatic diagnosis of various diseases from electronic medical records. We categorized our work into three classes 1) Rule-Based Methods, 2) Machine Learning Methods, and 3) Deep Learning Methods. These categories are further divided into sub-categories based on the proposed algorithm. In this review, we tried to cover almost all the latest and existing research of automatic diagnosis from electronic records. We presented the benefits, limitations, and future directions of various data-driven methods, dataset employed and focused disease. Moreover, we tried to establish a professional structure to familiarize with an up-to-date automatic disease diagnosis technique.

REFERENCES

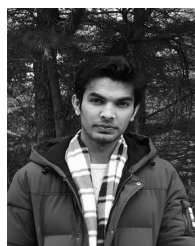
- [1] M. Stewart, *Patient-Centered Medicine: Transforming the Clinical Method*. Oxford, U.K.: Radcliffe Publishing, 2003.
- [2] J. Stausberg, D. Koch, J. Ingenerf, and M. Betzler, "Comparing paper-based with electronic patient records: Lessons learned during a study on diagnosis and procedure codes," *J. Amer. Med. Inform. Assoc.*, vol. 10, no. 5, pp. 470–477, Sep. 2003.
- [3] C. S. Kruse, R. Goswamy, Y. Raval, and S. Marawi, "Challenges and opportunities of big data in health care: A systematic review," *JMIR Med. Informat.*, vol. 4, no. 4, p. e38, Nov. 2016.
- [4] J. J. Firthous and M. M. Sathik, "Survey on using electronic medical records (EMR) to identify the health conditions of the patients," *J. Eng. Sci.*, vol. 11, no. 5, 2020.
- [5] G. Makoul, R. H. Curry, and P. C. Tang, "The use of electronic medical records: Communication patterns in outpatient encounters," *J. Amer. Med. Inform. Assoc.*, vol. 8, no. 6, pp. 610–615, Nov. 2001.
- [6] J. Chen, L. Sun, C. Guo, and Y. Xie, "A fusion framework to extract typical treatment patterns from electronic medical records," *Artif. Intell. Med.*, vol. 103, Mar. 2020, Art. no. 101782.
- [7] S. Ajami and T. Bagheri Tadi, "Barriers for adopting electronic health records (EHRs) by physicians," *Acta Inf. Medica*, vol. 21, no. 2, p. 129, 2013.
- [8] S. U. Rehman, S. Tu, Y. Huang, and Z. Yang, "Face recognition: A novel un-supervised convolutional neural network method," in *Proc. IEEE Int. Conf. Online Anal. Comput. Sci. (ICOACS)*, May 2016, pp. 139–144.
- [9] W. R. Hersh, "The electronic medical record: Promises and problems," *J. Amer. Soc. for Inf. Sci.*, vol. 46, no. 10, pp. 772–776, Dec. 1995.
- [10] N. Jenkins and R. Wilson, "The challenge of electronic health records (EHRs) design and implementation: Responses of health workers to drawing a 'big and rich picture' of a future EHR programme using animated tools," *J. Innov. Health Informat.*, vol. 15, no. 2, pp. 93–101, Jun. 2007.
- [11] M. Cifuentes, M. Davis, D. Fernald, R. Gunn, P. Dickinson, and D. J. Cohen, "Electronic health record challenges, workarounds, and solutions observed in practices integrating behavioral health and primary care," *J. Amer. Board Family Med.*, vol. 28, no. 1, pp. S63–S72, Sep. 2015.
- [12] M. N. Walsh, N. M. Albert, A. B. Curtis, M. Gheorghiadu, J. T. Heywood, Y. Liu, M. R. Mehra, C. M. O'Connor, D. Reynolds, C. W. Yancy, and G. C. Fonarow, "Lack of association between electronic health record systems and improvement in use of evidence-based heart failure therapies in outpatient cardiology practices," *Clin. Cardiol.*, vol. 35, no. 3, pp. 187–196, Mar. 2012.
- [13] S. ur Rehman, S. Tu, Y. Huang, and G. Liu, "CSFL: A novel unsupervised convolution neural network approach for visual pattern classification," *AI Commun.*, vol. 30, no. 5, pp. 311–324, Aug. 2017.
- [14] Z. Zeng, Y. Deng, X. Li, T. Naumann, and Y. Luo, "Natural language processing for EHR-based computational phenotyping," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 1, pp. 139–153, Jan. 2019.
- [15] S. Biafore, "Predictive solutions bring more power to decision makers," *Health Manage. Technol.*, vol. 20, no. 10, pp. 12–14, 1999.
- [16] G. W. Bassel, E. Glaab, J. Marquez, M. J. Holdsworth, and J. Bacardit, "Functional network construction in Arabidopsis using rule-based machine learning on large-scale data sets," *Plant Cell*, vol. 23, no. 9, pp. 3101–3116, 2011.

- [17] S. M. Weiss and N. Indurkha, "Rule-based machine learning methods for functional prediction," *J. Artif. Intell. Res.*, vol. 3, pp. 383–403, Dec. 1995.
- [18] C. Zhang and S. Zhang, *Association Rule Mining: Models and Algorithms*. Berlin, Germany: Springer-Verlag, 2002.
- [19] R. J. Urbanowicz and J. H. Moore, "Learning classifier systems: A complete introduction, review, and roadmap," *J. Artif. Evol. Appl.*, vol. 2009, pp. 1–25, 2009.
- [20] L. N. Castro, L. N. D. Castro, and J. Timmis, *Artificial Immune Systems: A New Computational Intelligence Approach*. Berlin, Germany: Springer, 2002.
- [21] H. F. El-Sofany and I. A. T. F. Taj-Eddin, "A cloud-based model for medical diagnosis using fuzzy logic concepts," in *Proc. Int. Conf. Innov. Trends Comput. Eng. (ITCE)*, Feb. 2019, pp. 162–167.
- [22] C. Shivade, P. Raghavan, E. Fosler-Lussier, P. J. Embi, N. Elhadad, S. B. Johnson, and A. M. Lai, "A review of approaches to identifying patient phenotype cohorts using electronic health records," *J. Amer. Med. Inform. Assoc.*, vol. 21, no. 2, pp. 221–230, Mar. 2014.
- [23] R. Zhang, S. Ma, L. Shanahan, J. Munroe, S. Horn, and S. Speedie, "Automatic methods to extract new york heart association classification from clinical notes," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2017, pp. 1296–1299.
- [24] L. S. Kumar and A. Padmapriya, "Rule based information extraction from electronic health records by forward-chaining," in *Article in Elsevier Ergonomics Book Series*. Amsterdam, The Netherlands, Aug. 2014.
- [25] P. B. Jensen, L. J. Jensen, and S. Brunak, "Mining electronic health records: Towards better research applications and clinical care," *Nature Rev. Genet.*, vol. 13, no. 6, pp. 395–405, 2012.
- [26] A. Jorge, V. M. Castro, A. Barnado, V. Gainer, C. Hong, T. Cai, T. Cai, R. Carroll, J. C. Denny, and L. Crofford, "Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms," in *Seminars in Arthritis and Rheumatism*. Amsterdam, The Netherlands: Elsevier, 2019.
- [27] E. Chazard, G. Ficheur, S. Bernonville, M. Luyckx, and R. Beuscart, "Data mining to generate adverse drug events detection rules," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 6, pp. 823–830, Nov. 2011.
- [28] H.-J. Dai, C.-Y. Chen, C.-Y. Wu, P.-T. Lai, R. T.-H. Tsai, and W.-L. Hsu, "Coreference resolution of medical concepts in discharge summaries by exploiting contextual information," *J. Amer. Med. Inform. Assoc.*, vol. 19, no. 5, pp. 888–896, Sep. 2012.
- [29] G. Hrovat, G. Stiglic, P. Kokol, and M. Ojsteršek, "Contrasting temporal trend discovery for large healthcare databases," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 251–257, Jan. 2014.
- [30] L. A. Petersen, S. Wright, S.-L.-T. Normand, and J. Daley, "Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database," *J. Gen. Internal Med.*, vol. 14, no. 9, pp. 555–558, Sep. 1999.
- [31] A. B. McCoy, D. F. Sittig, and A. Wright, "Comparison of association rule mining and crowdsourcing for automated generation of a problem-medication knowledge base," in *Proc. IEEE 2nd Int. Conf. Healthcare Informat., Imag. Syst. Biol.*, Sep. 2012, p. 125.
- [32] L. K. Wiley, J. D. Moretz, J. C. Denny, J. F. Peterson, and W. S. Bush, "Phenotyping adverse drug reactions: Statin-Related myotoxicity," *AMIA Summits Transl. Sci. Proc.*, vol. 2015, p. 466, Mar. 2015.
- [33] H. Ware, C. J. Mullett, and V. Jagannathan, "Natural language processing framework to assess clinical conditions," *J. Amer. Med. Inform. Assoc.*, vol. 16, no. 4, pp. 585–589, Jul. 2009.
- [34] A. N. Nguyen, M. J. Lawley, D. P. Hansen, R. V. Bowman, B. E. Clarke, E. E. Duhig, and S. Colquist, "Symbolic rule-based classification of lung cancer stages from free-text pathology reports," *J. Amer. Med. Inform. Assoc.*, vol. 17, no. 4, pp. 440–445, Jul. 2010.
- [35] H. Xu, Z. Fu, A. Shah, Y. Chen, N. B. Peterson, Q. Chen, S. Mani, M. A. Levy, Q. Dai, and J. C. Denny, "Extracting and integrating data from entire electronic health records for detecting colorectal cancer cases," in *Proc. AMIA Annu. Symp.*, 2011, p. 1564.
- [36] C. Friedman, "Medlee-a medical language extraction and encoding system," Queens College Cuny, Columbia Univ., New York, NY, USA, Tech. Rep. 938, 1995.
- [37] Q. Li, K. Melton, T. Lingren, E. S. Kirkendall, E. Hall, H. Zhai, Y. Ni, M. Kaiser, L. Stoutenborough, and I. Solti, "Phenotyping for patient safety: Algorithm development for electronic health record based automated adverse event and medical error detection in neonatal intensive care," *J. Amer. Med. Inform. Assoc.*, vol. 21, no. 5, pp. 776–784, Sep. 2014.
- [38] T. N. Raju, G. Suresh, and R. D. Higgins, "Patient safety in the context of neonatal intensive care: Research and educational opportunities," *Pediatric Res.*, vol. 70, no. 1, pp. 109–115, 2011.
- [39] M. Alodadi, "Radiology clinical notes mining using weighted association rules," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Aug. 2017, p. 325.
- [40] G. Szenasi, C. Lemnar, and I. Barbantan, "Concept extraction from medical documents a contextual approach," in *Proc. IEEE Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Sep. 2015, pp. 13–17.
- [41] C. Breischneider, S. Zillner, M. Hammon, P. Gass, and D. Sonntag, "Automatic extraction of breast cancer information from clinical reports," in *Proc. IEEE 30th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2017, pp. 213–218.
- [42] K. Haerian, H. Salmasian, and C. Friedman, "Methods for identifying suicide or suicidal ideation in EHRs," in *Proc. AMIA Annu. Symp.*, 2012, p. 1244.
- [43] B. C. Sauer, B. E. Jones, G. Globe, J. Leng, C.-C. Lu, T. He, C.-C. Teng, P. Sullivan, and Q. Zeng, "Performance of a natural language processing (NLP) tool to extract pulmonary function test (PFT) reports from structured and semistructured Veteran Affairs (VA) data," *eGEMS*, vol. 4, no. 1, pp. 1–18, 2016.
- [44] C.-S. Yu, Y.-J. Lin, C.-H. Lin, S.-Y. Lin, J. L. Wu, and S.-S. Chang, "Development of an online health care assessment for preventive medicine: A machine learning approach," *J. Med. Internet Res.*, vol. 22, no. 6, Jun. 2020, Art. no. e18585.
- [45] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 988–999, Sep. 1999.
- [46] X. Zhang, J. Xiao, and F. Gu, "Applying support vector machine to electronic health records for cancer classification," in *Proc. Spring Simul. Conf. (SpringSim)*, Apr. 2019, pp. 1–9.
- [47] T. S. Brisimi, T. Xu, T. Wang, W. Dai, W. G. Adams, and I. C. Paschalidis, "Predicting chronic disease hospitalizations from electronic health records: An interpretable classification approach," *Proc. IEEE*, vol. 106, no. 4, pp. 690–707, Apr. 2018.
- [48] X.-Y. Jin, D.-X. Pu, Y.-Z. Lan, and L.-J. Li, "Medical aided diagnosis using electronic medical records based on LDA and word vector model," in *Proc. 4th Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, Jul. 2017, pp. 443–445.
- [49] H. Alemzadeh and M. Devarakonda, "An NLP-based cognitive system for disease status identification in electronic health records," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Feb. 2017, pp. 89–92.
- [50] R. J. Carroll, A. E. Eyler, and J. C. Denny, "Naïve electronic health record phenotype identification for rheumatoid arthritis," in *Proc. AMIA Annu. Symp.*, 2011, p. 189.
- [51] Q. T. Zeng, S. Goryachev, S. Weiss, M. Sordo, S. N. Murphy, and R. Lazarus, "Extracting principal diagnosis, co-morbidity and smoking status for asthma research: Evaluation of a natural language processing system," *BMC Med. Informat. Decis. Making*, vol. 6, no. 1, p. 30, Dec. 2006.
- [52] C. A. Bejan, F. Xia, L. Vanderwende, M. M. Wurfel, and M. Yetisgen-Yildiz, "Pneumonia identification using statistical feature selection," *J. Amer. Med. Inform. Assoc.*, vol. 19, no. 5, pp. 817–823, Sep. 2012.
- [53] S. U. Rehman, S. Tu, Y. Huang, and O. U. Rehman, "A benchmark dataset and learning high-level semantic embeddings of multimedia for cross-media retrieval," *IEEE Access*, vol. 6, pp. 67176–67188, 2018.
- [54] V. Garla, C. Taylor, and C. Brandt, "Semi-supervised clinical text classification with Laplacian SVMs: An application to cancer case management," *J. Biomed. Informat.*, vol. 46, no. 5, pp. 869–875, Oct. 2013.
- [55] Z. Zeng, X. Li, S. Espino, A. Roy, K. Kitsch, S. Clare, S. Khan, and Y. Luo, "Contralateral breast cancer event detection using nature language processing," in *Proc. AMIA Annu. Symp.*, 2017, p. 1885.
- [56] Y. Chen, R. J. Carroll, E. R. M. Hinz, A. Shah, A. E. Eyler, J. C. Denny, and H. Xu, "Applying active learning to high-throughput phenotyping algorithms for electronic health records data," *J. Amer. Med. Inform. Assoc.*, vol. 20, no. e2, pp. e253–e259, Dec. 2013.
- [57] W.-Q. Wei, C. Tao, G. Jiang, and C. G. Chute, "A high throughput semantic concept frequency based approach for patient identification: A case study using type 2 diabetes mellitus clinical notes," in *Proc. AMIA Annu. Symp.*, 2010, p. 857.
- [58] K. M. Z. Hasan, S. Datta, M. Z. Hasan, and N. Zahan, "Automated prediction of heart disease patients using sparse discriminant analysis," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2019, pp. 1–6.

- [59] I. Rish, "An empirical study of the naive Bayes classifier," in *Proc. Workshop Empirical Methods Artif. Intell. (IJCAI)* vol. 3. New York, NY, USA: IBM, 2001, pp. 41–46.
- [60] R. Rana and J. Pruthi, "Heart disease prediction using Naïve Bayes classification in data mining," *IJSRD-Int. J. Sci. Res. Develop.*, vol. 2, no. 05, pp. 2321–0613, 2014.
- [61] K. M. Al-Aidaroo, A. A. Bakar, and Z. Othman, "Medical data classification with naive bayes approach," *Inf. Technol. J.*, vol. 11, no. 9, pp. 1166–1174, Sep. 2012.
- [62] I. Kononenko, "Inductive and Bayesian learning in medical diagnosis," *Appl. Artif. Intell.*, vol. 7, no. 4, pp. 317–337, Oct. 1993.
- [63] K.-J. Wang, B. Makond, and K.-M. Wang, "Modeling and predicting the occurrence of brain metastasis from lung cancer by Bayesian network: A case study of taiwan," *Comput. Biol. Med.*, vol. 47, pp. 147–160, Apr. 2014.
- [64] J. Kazmierska and J. Malicki, "Application of the Naïve Bayesian Classifier to optimize treatment decisions," *Radiotherapy Oncol.*, vol. 86, no. 2, pp. 211–216, 2008.
- [65] S. Sakai, K. Kobayashi, J. Nakamura, S. Toyabe, and K. Akazawa, "Accuracy in the diagnostic prediction of acute appendicitis based on the Bayesian network model," *Methods Inf. Med.*, vol. 46, no. 06, pp. 723–726, 2007.
- [66] K. Vembandasamy, R. Sasipriya, and E. Deepa, "Heart diseases detection using Naïve Bayes algorithm," *Int. J. Innov. Sci., Eng. Technol.*, vol. 2, no. 9, pp. 441–444, 2015.
- [67] W. Wei, S. Visweswaran, and G. F. Cooper, "The application of naive bayes model averaging to predict Alzheimer's disease from genome-wide data," *J. Amer. Med. Inform. Assoc.*, vol. 18, no. 4, pp. 370–375, Jul. 2011.
- [68] M. Anbarasi, E. Anupriya, and N. C. S. N. Iyengar, "Enhanced prediction of heart disease with feature subset selection using genetic algorithm," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 10, pp. 5370–5376, 2010.
- [69] E. Aličković and A. Subasi, "Data mining techniques for medical data classification," in *Proc. Int. Arab Conf. Inf. Technol. (ACIT)*, 2011, pp. 1–5.
- [70] S. Anto and S. Chandramathi, "Supervised machine learning approaches for medical data set classification-a review," *Int. J. Comput. Sci. Technol.*, vol. 2, no. 4, pp. 234–240, 2011.
- [71] L. Tanner, M. Schreiber, J. G. H. Low, A. Ong, T. Tolfvenstam, Y. L. Lai, L. C. Ng, Y. S. Leo, L. Thi Puong, S. G. Vasudevan, C. P. Simmons, M. L. Hibberd, and E. E. Ooi, "Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness," *PLoS Neglected Tropical Diseases*, vol. 2, no. 3, p. e196, Mar. 2008.
- [72] V. Sabarinathan and V. Sugumaran, "Diagnosis of heart disease using decision tree," *Int. J. Res. Comput. Appl. Inf. Technol.*, vol. 2, no. 6, pp. 74–79, 2014.
- [73] W.-J. Kuo, R.-F. Chang, D.-R. Chen, and C. C. Lee, "Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images," *Breast Cancer Res. Treatment*, vol. 66, no. 1, pp. 51–57, Mar. 2001.
- [74] A. Tyagi and P. Singh, "Asthma diagnosis and level of control using decision tree and fuzzy system," *Int. J. Biomed. Eng. Technol.*, vol. 16, no. 2, pp. 169–181, 2014.
- [75] M. Shouman, T. Turner, and R. Stocker, "Using decision tree for diagnosing heart disease patients," in *Proc. 9th Australas. Data Mining Conf.*, vol. 121, 2011, pp. 23–30.
- [76] V.-S. Sowmien, V. Sugumaran, P. Karthikeyan, and T.-R. Vijayaram, "Diagnosis of hepatitis using decision tree algorithm," *Int. J. Eng. Technol.*, vol. 8, no. 3, pp. 1414–1419, 2016.
- [77] A. Lungu, A. J. Swift, D. Capener, D. Kiely, R. Hose, and J. M. Wild, "Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis," *Pulmonary Circulat.*, vol. 6, no. 2, pp. 181–190, 2016.
- [78] Y. Sun and D. Zhang, "Diagnosis and analysis of diabetic retinopathy based on electronic health records," *IEEE Access*, vol. 7, pp. 86115–86120, 2019.
- [79] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, "A review of relational machine learning for knowledge graphs," *Proc. IEEE*, vol. 104, no. 1, pp. 11–33, Jan. 2016.
- [80] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.
- [81] Y.-W. Chen and L. C. Jain, *Deep Learning in Healthcare*. Berlin, Germany: Springer, 2020.
- [82] K. Muhammad, S. Khan, J. D. Ser, and V. H. C. de Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 30, 2020, doi: 10.1109/TNNLS.2020.2995800.
- [83] M. W. Nadeem, M. A. A. Ghamdi, M. Hussain, M. A. Khan, K. M. Khan, S. H. Almotiri, and S. A. Butt, "Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges," *Brain Sci.*, vol. 10, no. 2, p. 118, Feb. 2020.
- [84] L. Franz, Y. Raj Shrestha, and B. Paudel, "A deep learning pipeline for patient diagnosis prediction using electronic health records," 2020, *arXiv:2006.16926*. [Online]. Available: <http://arxiv.org/abs/2006.16926>
- [85] M. Usama, B. Ahmad, W. Xiao, M. S. Hossain, and G. Muhammad, "Self-attention based recurrent convolutional neural network for disease prediction using healthcare data," *Comput. Methods Programs Biomed.*, vol. 190, Jul. 2020, Art. no. 105191.
- [86] X. Shi, Y. Hu, Y. Zhang, W. Li, Y. Hao, A. Alelaiwi, S. M. M. Rahman, and M. Shamim Hossain, "Multiple disease risk assessment with uniform model based on medical clinical notes," *IEEE Access*, vol. 4, pp. 7074–7083, 2016.
- [87] Y. Sun, "The neural network of one-dimensional Convolution—An example of the diagnosis of diabetic retinopathy," *IEEE Access*, vol. 7, pp. 69657–69666, 2019.
- [88] J. Mullenbach, S. Wiegrefe, J. Duke, J. Sun, and J. Eisenstein, "Explainable prediction of medical codes from clinical text," 2018, *arXiv:1802.05695*. [Online]. Available: <http://arxiv.org/abs/1802.05695>
- [89] R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley, "Deep patient: An unsupervised representation to predict the future of patients from the electronic health records," *Sci. Rep.*, vol. 6, no. 1, p. 26094, May 2016.
- [90] E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," *J. Amer. Med. Informat. Assoc.*, vol. 24, no. 2, pp. 361–370, Mar. 2017.
- [91] Y. Cheng, F. Wang, P. Zhang, and J. Hu, "Risk prediction with electronic health records: A deep learning approach," in *Proc. SIAM Int. Conf. Data Mining*, Jun. 2016, pp. 432–440.
- [92] Z. Zhu, C. Yin, B. Qian, Y. Cheng, J. Wei, and F. Wang, "Measuring patient similarities via a deep architecture with medical concept embedding," in *Proc. IEEE 16th Int. Conf. Data Mining (ICDM)*, Dec. 2016, pp. 749–758.
- [93] S. Mehrabi, S. Sohn, D. Li, J. J. Pankratz, T. Therneau, J. L. St. Sauver, H. Liu, and M. Palakal, "Temporal pattern and association discovery of diagnosis codes using deep learning," in *Proc. Int. Conf. Healthcare Informat.*, Oct. 2015, pp. 408–416.
- [94] L. Nie, M. Wang, L. Zhang, S. Yan, B. Zhang, and T.-S. Chua, "Disease inference from health-related questions via sparse deep learning," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 8, pp. 2107–2119, Aug. 2015.
- [95] S. Rehman, S. Tu, O. Rehman, Y. Huang, C. Magurawalage, and C.-C. Chang, "Optimization of CNN through novel training strategy for visual classification problems," *Entropy*, vol. 20, no. 4, p. 290, Apr. 2018.
- [96] S. U. Rehman, S. Tu, M. Waqas, Y. Huang, O. U. Rehman, B. Ahmad, and S. Ahmad, "Unsupervised pre-trained filter learning approach for efficient convolution neural network," *Neurocomputing*, vol. 365, pp. 171–190, Nov. 2019.
- [97] W. Geng, X. Qin, Z. Wang, Q. Kong, Z. Tang, and L. Jiang, "Model-based reasoning methods for diagnosis in integrative medicine based on electronic medical records and natural language processing," *medRxiv*, to be published. [Online]. Available: <https://www.medrxiv.org/content/early/2020/07/18/2020.07.12.20151746>, doi: 10.1101/2020.07.12.20151746.
- [98] T. Baumel, J. Nassour-Kassis, R. Cohen, M. Elhadad, and N. Elhadad, "Multi-label classification of patient notes a case study on ICD code assignment," 2018, *arXiv:1709.09587*. [Online]. Available: <https://arxiv.org/abs/1709.09587>
- [99] J. M. Brown, J. P. Campbell, A. Beers, K. Chang, S. Ostmo, R. V. P. Chan, J. Dy, D. Erdogmus, S. Ioannidis, J. Kalpathy-Cramer, M. F. Chiang, f. the Imaging, and I. in Retinopathy of Prematurity (i-ROP) Research Consortium, "Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks," *JAMA Ophthalmol.*, vol. 136, no. 7, p. 803, Jul. 2018.
- [100] B. Jin, C. Che, Z. Liu, S. Zhang, X. Yin, and X. Wei, "Predicting the risk of heart failure with EHR sequential data modeling," *IEEE Access*, vol. 6, pp. 9256–9261, 2018.
- [101] J. Shi, X. Fan, J. Wu, J. Chen, and W. Chen, "DeepDiagnosis: DNN-based diagnosis prediction from pediatric big healthcare data," in *Proc. 6th Int. Conf. Adv. Cloud Big Data (CBD)*, Aug. 2018, pp. 287–292.

- [102] W. W. Wang, H. Li, L. Cui, X. Hong, and Z. Yan, "Predicting clinical visits using recurrent neural networks and demographic information," in *Proc. IEEE 22nd Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2018, pp. 353–358.
- [103] S. Wu, S. Liu, S. Sohn, S. Moon, C.-I. Wi, Y. Juhn, and H. Liu, "Modeling asynchronous event sequences with RNNs," *J. Biomed. Informat.*, vol. 83, pp. 167–177, Jul. 2018.
- [104] J. Chu, W. Dong, K. He, H. Duan, and Z. Huang, "Using neural attention networks to detect adverse medical events from electronic health records," *J. Biomed. Informat.*, vol. 87, pp. 118–130, Nov. 2018.
- [105] Y. Luo, "Recurrent neural networks for classifying relations in clinical notes," *J. Biomed. Informat.*, vol. 72, pp. 85–95, Aug. 2017.
- [106] X. Dong, S. Chowdhury, L. Qian, Y. Guan, J. Yang, and Q. Yu, "Transfer bi-directional LSTM RNN for named entity recognition in chinese electronic medical records," in *Proc. IEEE 19th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2017, pp. 1–4.
- [107] C. Esteban, O. Staeck, S. Baier, Y. Yang, and V. Tresp, "Predicting clinical events by combining static and dynamic information using recurrent neural networks," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Oct. 2016, pp. 93–101.
- [108] Y. Hao, M. Usama, J. Yang, M. S. Hossain, and A. Ghoneim, "Recurrent convolutional neural network based multimodal disease risk prediction," *Future Gener. Comput. Syst.*, vol. 92, pp. 76–83, Mar. 2019.
- [109] S. Farzi, S. Kianian, and I. Rastkhadive, "Diagnosis of attention deficit hyperactivity disorder using deep belief network based on greedy approach," in *Proc. 5th Int. Symp. Comput. Bus. Intell. (ISCBI)*, Aug. 2017, pp. 96–99.
- [110] M. Faturrahman, I. Wasito, N. Hanifah, and R. Mufidah, "Structural MRI classification for Alzheimer's disease detection using deep belief network," in *Proc. 11th Int. Conf. Inf. Commun. Technol. Syst. (ICTS)*, Oct. 2017, pp. 37–42.
- [111] A. H. Al-Fatlawi, M. H. Jabardi, and S. H. Ling, "Efficient diagnosis system for Parkinson's disease using deep belief network," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2016, pp. 1324–1330.
- [112] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [113] Z. Zhao, L. Jiao, J. Zhao, J. Gu, and J. Zhao, "Discriminant deep belief network for high-resolution SAR image classification," *Pattern Recognit.*, vol. 61, pp. 686–701, Jan. 2017.
- [114] X.-L. Hao and M. Tian, "Deep belief network based on double weber local descriptor in micro-expression recognition," in *Advanced Multimedia and Ubiquitous Engineering*. Berlin, Germany: Springer, 2017, pp. 419–425.
- [115] X. Gao, C. Shang, Y. Jiang, D. Huang, and T. Chen, "Refinery scheduling with varying crude: A deep belief network classification and multimodel approach," *AICHE J.*, vol. 60, no. 7, pp. 2525–2532, Jul. 2014.
- [116] C. Shang, F. Yang, D. Huang, and W. Lyu, "Data-driven soft sensor development based on deep learning technique," *J. Process Control*, vol. 24, no. 3, pp. 223–233, Mar. 2014.
- [117] Y. Lyu, J. Chen, and Z. Song, "Image-based process monitoring using deep belief networks," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 115–120, 2018.
- [118] Z. Geng, Z. Li, and Y. Han, "A new deep belief network based on RBM with glial chains," *Inf. Sci.*, vols. 463–464, pp. 294–306, Oct. 2018.
- [119] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," 2013, *arXiv:1312.6114*. [Online]. Available: <http://arxiv.org/abs/1312.6114>
- [120] A. B. L. Larsen and S. K. Sønderby. (Nov. 13, 2015). *Generating Faces With Torch*. Accessed: Apr. 27, 2017. [Online]. Available: <http://arxiv.org/abs/1506.05751>
- [121] P. Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York, NY, USA: Basic Books, 2015.
- [122] K. Adem, S. Kilicarslan, and O. Cömert, "Classification and diagnosis of cervical cancer with softmax classification with stacked autoencoder," *Expert Syst. Appl.*, vol. 115, pp. 557–564, Jan. 2019.
- [123] T. Zhang, A. H. Sodhro, Z. Luo, N. Zahid, M. W. Nawaz, S. Pirbhulal, and M. Muzammal, "A joint deep learning and Internet of medical things driven framework for elderly patients," *IEEE Access*, vol. 8, pp. 75822–75832, 2020.
- [124] Y. Deng, A. Sander, L. Faulstich, and K. Denecke, "Towards automatic encoding of medical procedures using convolutional neural networks and autoencoders," *Artif. Intell. Med.*, vol. 93, pp. 29–42, Jan. 2019.
- [125] U. Hwang, S. Choi, H.-B. Lee, and S. Yoon, "Adversarial training for disease prediction from electronic health records with missing data," 2017, *arXiv:1711.04126*. [Online]. Available: <http://arxiv.org/abs/1711.04126>
- [126] B. Abraham and M. S. Nair, "Computer-aided diagnosis of clinically significant prostate cancer from MRI images using sparse autoencoder and random forest classifier," *Biocybern. Biomed. Eng.*, vol. 38, no. 3, pp. 733–744, 2018.
- [127] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica, "Spark: Cluster computing with working sets," *HotCloud*, vol. 10, nos. 10–10, p. 95, Jun. 2010.
- [128] T. White, *Hadoop: The Definitive Guide*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2012.
- [129] S. U. Rehman, Y. Huang, S. Tu, and O. U. Rehman, "Facebook5k: A novel evaluation resource dataset for cross-media search," in *Proc. Int. Conf. Cloud Comput. Secur.* Haikou, China: Springer, 2018, pp. 512–524.
- [130] V. Pabbi, "Fuzzy expert system for medical diagnosis," *Int. J. Sci. Res. Publications*, vol. 5, no. 1, pp. 1–7, 2015.
- [131] R. Main and K. Oldham, "Increasing the acceptance of knowledge-based systems," in *Proc. 13th Int. MATADOR Conf.* Springer, 1993, pp. 303–313.
- [132] K. A. Spackman and D. Connelly, "Knowledge-based systems in laboratory medicine and pathology. A review and survey of the field," *Arch. Pathol. Lab. Med.*, vol. 111, no. 2, pp. 116–119, 1987.
- [133] F. S. Roque, P. B. Jensen, H. Schmock, M. Dalgaard, M. Andreatta, T. Hansen, K. Søbey, S. Bredkjær, A. Juul, T. Werge, L. J. Jensen, and S. Brunak, "Using electronic patient records to discover disease correlations and stratify patient cohorts," *PLoS Comput. Biol.*, vol. 7, no. 8, Aug. 2011, Art. no. e1002141.
- [134] G. K. Savova, K. C. Kipper-Schuler, J. F. Hurdle, and S. M. Meystre, "Extracting information from textual documents in the electronic health record: A review of recent research," *Yearbook Med. Informat.*, vol. 17, no. 1, pp. 128–144, 2008.
- [135] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 715–718.
- [136] M. D. Abramoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, and M. Niemeijer, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Invest. Ophthalmol. Vis. Sci.*, vol. 57, no. 13, pp. 5200–5206, 2016.
- [137] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, and J. Cuadros, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [138] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sanchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [139] P. Prentašić and S. Lončarić, "Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion," *Comput. Methods Programs Biomed.*, vol. 137, pp. 281–292, Dec. 2016.
- [140] D. E. Worrall, C. M. Wilson, and G. J. Brostow, "Automated retinopathy of prematurity case detection with convolutional neural networks," in *Deep Learning and Data Labeling for Medical Applications*. Athens, Greece: Springer, 2016, pp. 68–76.
- [141] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, "Detection of age-related macular degeneration via deep learning," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 184–188.
- [142] D. Bychkov, R. Turkki, C. Haglund, N. Linder, and J. Lundin, "Deep learning for tissue microarray image-based outcome prediction in patients with colorectal cancer," *Proc. SPIE*, vol. 9791, Mar. 2016, Art. no. 979115.
- [143] A. Cruz-Roa, A. Basavanahally, F. González, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Proc. SPIE*, vol. 9041, Mar. 2014, Art. no. 904103.
- [144] G. Litjens, C. I. Sánchez, N. Timofeeva, M. Hermesen, I. Nagtegaal, I. Kovacs, C. Hulsbergen-van de Kaa, P. Bult, B. van Ginneken, and J. van der Laak, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Sci. Rep.*, vol. 6, no. 1, p. 26286, Sep. 2016.

- [145] H. Rezaeilouyeh, A. Mollahosseini, and M. H. Mahoor, "Microscopic medical image classification framework via deep learning and shearlet transform," *J. Med. Imag.*, vol. 3, no. 4, Nov. 2016, Art. no. 044501.
- [146] J. A. Quinn, R. Nakasi, P. K. Mugagga, P. Byanyima, W. Lubega, and A. Andama, "Deep convolutional neural networks for microscopy-based point of care diagnostics," in *Proc. Mach. Learn. Healthcare Conf.*, 2016, pp. 271–281.
- [147] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, "Deep learning for identifying metastatic breast cancer," 2016, *arXiv:1606.05718*. [Online]. Available: <http://arxiv.org/abs/1606.05718>
- [148] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I.-C. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 1626–1630.
- [149] M. G. Ertosun and D. L. Rubin, "Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks," in *Proc. AMIA Annu. Symp.*, 2015, p. 1899.
- [150] H. Chang, J. Han, C. Zhong, A. M. Snijders, and J.-H. Mao, "Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1182–1194, May 2018.
- [151] E. Kim, M. Corte-Real, and Z. Baloch, "A deep semantic mobile application for thyroid cytopathology," *Proc. SPIE*, vol. 9789, Apr. 2016, Art. no. 97890A.
- [152] C. Mao, L. Yao, Y. Pan, Y. Luo, and Z. Zeng, "Deep generative classifiers for thoracic disease diagnosis with chest X-ray images," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2018, pp. 1209–1214.
- [153] T. M. Tuan, N. T. Duc, P. Van Hai, and L. H. Son, "Dental diagnosis from X-ray images using fuzzy rule-based systems," *Int. J. Fuzzy Syst. Appl.*, vol. 6, no. 1, pp. 1–16, Jan. 2017.
- [154] S. A. Ahmad, M. N. Taib, N. E. A. Khalid, and H. Taib, "An analysis of image enhancement techniques for dental X-ray image interpretation," *Int. J. Mach. Learn. Comput.*, vol. 2, no. 3, p. 292, 2012.
- [155] K. A. Jobst, L. P. D. Barnettson, and B. J. Shepstone, "Accurate prediction of histologically confirmed Alzheimer's disease and the differential diagnosis of dementia: The use of NINCDS-ADRDA and DSM-III-R criteria, SPECT, X-ray CT, and apo e4 in medial temporal lobe dementias," *Int. Psychogeriatrics*, vol. 9, no. S1, pp. 191–222, Dec. 1997.
- [156] U. J. Schoepf and P. Costello, "CT angiography for diagnosis of pulmonary embolism: State of the art," *Radiology*, vol. 230, no. 2, pp. 329–337, Feb. 2004.
- [157] D. A. Lynch, T. Hay, J. N. Jr, V. Divgi, and L. Fan, "Pediatric diffuse lung disease: Diagnosis and classification using high-resolution CT," *Amer. J. Roentgenol.*, vol. 173, no. 3, pp. 713–718, 1999.
- [158] E.-L. Chen, P.-C. Chung, C.-L. Chen, H.-M. Tsai, and C.-I. Chang, "An automatic diagnostic system for CT liver image classification," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 783–794, Jun. 1998.
- [159] P. J. Haug, L. Christensen, M. Gundersen, B. Clemons, S. Koehler, and K. Bauer, "A natural language parsing system for encoding admitting diagnoses," in *Proc. AMIA Annu. Fall Symp.*, 1997, p. 814.
- [160] D. Nadeau and S. Sekine, "A survey of named entity recognition and classification," *Linguistic Investigations*, vol. 30, no. 1, pp. 3–26, Jan. 2007.
- [161] H. Xu, S. P. Stenner, S. Doan, K. B. Johnson, L. R. Waitman, and J. C. Denny, "MedEx: A medication information extraction system for clinical narratives," *J. Amer. Med. Inform. Assoc.*, vol. 17, no. 1, pp. 19–24, Jan. 2010.
- [162] S. Tu, S. ur Rehman, M. Waqas, O. U. Rehman, Z. Yang, B. Ahmad, Z. Halim, and W. Zhao, "Optimisation-based training of evolutionary convolution neural network for visual classification applications," *IET Comput. Vis.*, vol. 14, no. 5, pp. 259–267, Aug. 2020.
- [163] T. Sajana and M. Narasingarao, "Machine learning techniques for malaria disease diagnosis—a review," *J. Adv. Res. Dyn. Control Syst.*, vol. 9, no. 6, pp. 349–369, 2017.
- [164] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology identification using deep feature selection with non-medical training," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 3, pp. 259–263, May 2018.
- [165] M. Cicero, A. Bilbily, E. Colak, T. Dowdell, B. Gray, K. Perampaladas, and J. Barfett, "Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs," *Investigative Radiol.*, vol. 52, no. 5, pp. 281–287, May 2017.
- [166] Z. Gao, L. Wang, L. Zhou, and J. Zhang, "HEp-2 cell image classification with deep convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 2, pp. 416–428, Mar. 2017.
- [167] X.-H. Han, J. Lei, and Y.-W. Chen, "HEp-2 cell classification using K-support spatial pooling in deep CNNs," in *Deep Learning and Data Labeling for Medical Applications*. Springer, 2016, pp. 3–11.
- [168] S. U. Rehman, Y. Huang, S. Tu, and B. Ahmad, "Learning a semantic space for modeling images, tags and feelings in cross-media search," in *Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining*, Macau, China: Springer, 2019, pp. 65–76.
- [169] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of MR brain images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Nagoya, Japan: Springer, 2013, pp. 649–656.
- [170] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, and N. Komodakis, "A deep metric for multimodal registration," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Athens, Greece: Springer, 2016, pp. 10–18.
- [171] X. Cheng, L. Zhang, and Y. Zheng, "Deep similarity learning for multimodal medical images," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 3, pp. 248–252, May 2018.
- [172] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Athens, Greece: Springer, 2016, pp. 132–139.
- [173] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 698–701.
- [174] A. Imran, J. Li, Y. Pei, J.-J. Yang, and Q. Wang, "Comparative analysis of vessel segmentation techniques in retinal images," *IEEE Access*, vol. 7, pp. 114862–114887, 2019.
- [175] A. Imran, J. Li, Y. Pei, F. Akhtar, J.-J. Yang, and Q. Wang, "Cataract detection and grading with retinal images using SOM-RBF neural network," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2019, pp. 2626–2632.



JAHANZAIB LATIF received the bachelor's degree from Bahauddin Zakariya University, Pakistan, in 2013, and the master's degree in computer science from the COMSATS Institute of Information Technology, Pakistan, in 2016. He is currently pursuing the Ph.D. degree in computer science with the Beijing University of Technology, Beijing, China. His research interests include machine learning, medical image analysis, and neural networks.



CHUANGBAI XIAO received the Ph.D. degree from Tsinghua University, in 1995. He was with the Beijing University of Aeronautics and Astronautics as a Postdoctoral Researcher, from 1995 to 1996. He was also with the Beijing Post Administration Bureau, Beijing Post Science Research Institute, from 1995 to 2001. Since 2001, he has been a Professor with the Faculty of Information Technology, Beijing University of Technology, China. His research interests include signal processing, pattern recognition and intelligent information systems, video information processing, and multimedia communication.



SHANSHAN TU (Member, IEEE) received the Ph.D. degree from the Computer Science Department, Beijing University of Posts and Telecommunications, in 2014. He visited the University of Essex for joint doctoral training, from 2013 to 2014. He was with the Department of Electronic Engineering, Tsinghua University, as a Postdoctoral Researcher, from 2014 to 2016. He is currently an Assistant Professor with the Faculty of Information Technology, Beijing University of

Technology, China. His research interests include cloud computing security and deep learning.



SADAQAT UR REHMAN (Member, IEEE) received the Ph.D. degree from the Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing, China, in 2019. He was with the Sarhad University of Science and Information Technology, Peshawar, Pakistan, as a Lecturer, from 2012 to 2015. He is currently an Assistant Professor with the Department of Computer Science, Namal Institute–Associated College of the University of

Bradford, U.K. He is also a Visiting Lecturer at the Department of Computer Science, Beijing University of Technology, China. His current research interests include data science, including deep learning, machine learning, and optimization techniques. He has a range of publications in these fields in the conferences and journals of repute. He is a Co-Chair, a TPC Member, and a Reviewer for prestigious international conferences and journals, including Smarttech 2020, CSAE 2018, 2019, and 2020, and IEEE Access.



AZHAR IMRAN (Graduate Student Member, IEEE) received the B.S. degree in software engineering and the M.S. degree in computer science from the University of Sargodha, Pakistan, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Beijing University of Technology, Beijing, China. From 2012 to 2017, he was a Lecturer with the Department of Computer Science, University of Sargodha. He has published many articles in reputed journals and

conferences. His research interests include machine learning, image processing, medical imaging, and data mining. His awards and honors include the China Scholarship Council (CSC), China, and the Star Contribution Award for best researcher 2018 from the Beijing University of Technology.



ANAS BILAL received the B.S. degree in telecommunication and networks from Iqra University, Pakistan, in 2013, and the M.S. degree in electrical and electronic systems from the University of Lahore, Pakistan, in 2016. He is currently pursuing the Ph.D. degree with the School of Information Technology, Beijing University of Technology. His research interests include neural networks, medical image analysis, machine learning, sentiment analysis, pattern recognition, signal and

image processing, geometrical channel modeling, channel estimation, and mobile communications.

...