JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Systematically Investigating the Role of Deep Learning in Diabetes

harshita Rathi, Dr. Ashok kumar Behera
Student, associate Professor
Bhilai Institute of technology

Abstract: - *Around 470 million people in the globe have diabetes, a chronic metabolic condition. Digital health has been developed with the goal of bettering the care of diabetic patients. Recent years' broad usage has produced a lot of data that may be utilised to inform next initiatives to end this chronic illness. Deep learning, a relatively new kind of machine learning with intriguing prospective applications, is one approach that has benefited from this transition. In this study, we assess the status of deep learning applications currently used in the study of diabetes. The phases of diabetes treatment that this approach is most often employed in include diagnosis, glucose control, and complication identification, according to a study of the literature. We have emphasised the most important information from the 40 original research publications we selected based on our search about the learning models used, the development process, the major results, and the baseline methodologies for performance measurement.*

According to the reviewed literature, it is now feasible to accomplish numerous tasks related to diabetes with state-of-the-art accuracy by using deep learning frameworks and algorithms, which perform better than more conventional machine learning approaches. In the meanwhile, we draw attention to a number of gaps in the existing research, such as a dearth of readily available data and uncertainty in model interpretation. Rapid advancements in deep learning and an abundance of data suggest that these issues may soon be resolved, enabling further use of this technology in therapeutic settings.

Keyword's: - **Diabetes, deep learning, deep neural networks, glucose management, diabetic complications, artificial intelligence**

INTRODUCTION: -

Chronic metabolic illnesses characterised by inadequate insulin production or diminished insulin activity are collectively referred to as diabetes. Due to the disease's complicated origins, the International Diabetes Federation predicts that in 2019, 463 million individuals throughout the world will have diabetes. However, experts believe that as many as half of these patients will go undetected. The 95% confidence interval for this estimate is 369-601 million. In the next ten years, diabetes is projected to become epidemic. Therefore, particularly in low- and middle-income countries, the prevention and treatment of diabetes has been a significant drain on national economies, healthcare systems, and individual medical expenses [1].

Clinical classifications of diabetes may be broken down into three groups according to their etiopathology: type 1 diabetes (T1D), type 2 diabetes (T2D), and gestational diabetes mellitus (GDM). Adult-onset diabetes from latent autoimmunity and juvenile-onset diabetes that improves with age are two more subcategories with a different an etiology. When the pancreatic cells that produce insulin are attacked and destroyed by the immune system, the result is type 1 diabetes [2, 3]. People with T1D need exogenous insulin therapy due to inadequate endocrine insulin synthesis. Insulin resistance or inadequate insulin production leads to the development of type 2 diabetes (T2D), the most prevalent form of the disease. Diagnosis of gestational diabetes mellitus (GDM) may need dietary adjustments and, in certain cases, the administration of exogenous insulin to ensure the health of the growing child. Early diagnosis and categorization of diabetes are challenging due to increased variability and a lack of continuous surveillance [2].

Most diabetics who need exogenous insulin stick to a schedule of MDI using an insulin pen or an insulin pump (continuous subcutaneous insulin infusion, or CSII), all while keeping a constant check on their blood sugar levels using a metre. It is crucial for diabetics to keep their blood glucose (BG) levels within a safe range at all times. Neuropathy, nephropathy, retinopathy, stroke, cardiovascular disease, and peripheral vascular disease are all examples of microvascular and macrovascular problems that may result from either hyperglycemia or hypoglycemia. However, because to the vast range of everyday activities (eating, exercising, drinking alcohol, being unwell, and stress) that may effect BG levels, it may be difficult for diabetics to maintain stable BG levels. Self-management practises, such as prompt blood glucose (BG) monitoring, hormone supply, and adherence to prescribed lifestyle recommendations, are thus crucial, but they need interdisciplinary clinical practise expertise, particularly for children and adolescents. The issue of establishing an effective treatment technique for an individual patient is further complicated by the high degrees of inter- and intra-population variation in the glucose kinetics process and pharmacokinetics. [4]

The artificial pancreas (AP), also known as closed-loop hormone delivery systems and continuous glucose monitoring (CGM) systems [4], has been the subject of much research in recent years. The research' objectives are to achieve automated glycemic control and reduce the workload associated with managing glucose levels. An AP system uses CSII by means of an insulin pump, closed-loop control algorithm, and continuous glucose monitoring (CGM). Several T1D support organisations advocate it since it has been shown to significantly improve glycemic control [5]. With the development of smart pens and smart metres that can communicate wirelessly with a smartphone, basal-bolus insulin therapy with a capillary blood glucose metre and MDI continues to be a cost-effective alternative for managing diabetes. When it comes to insulin distribution, the AP is state-of-the-art.

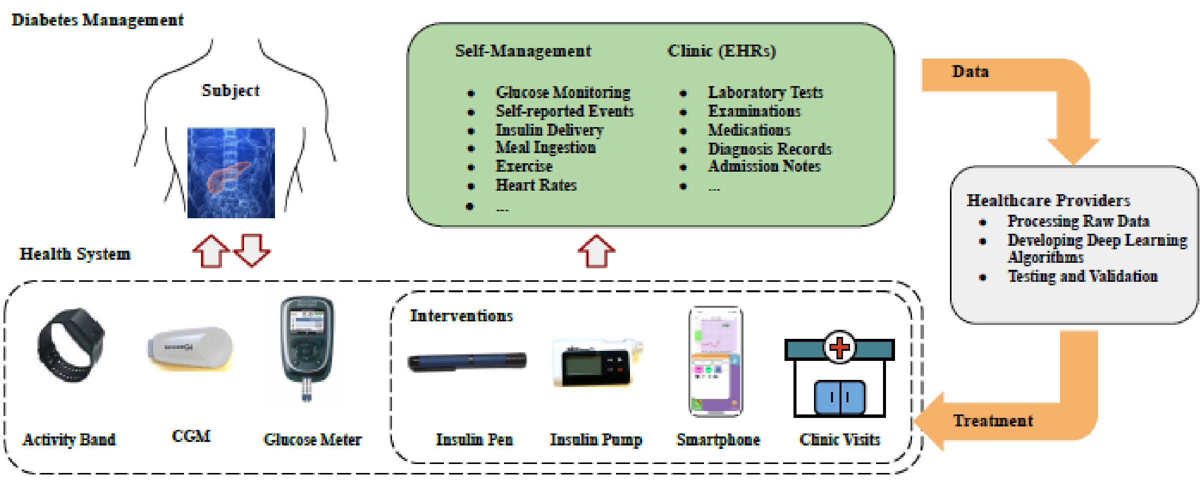

Figure 1: The illustration of diabetes management. The data is processed by healthcare providers to develop deep learning algorithms for novel therapies.

Improvements in smartphone apps and CGM integration for diabetes management have made it possible for users to keep tabs on daily activities and make educated treatment choices. Using resistance bands for BG management [6]. Wearables, digital tools, and electronic health records (EHRs) generate a great deal of information. The use of AI in this context may result in better care for diabetics.

It's possible that high-dimensional, sparse, multi-source medical datasets are underutilised in clinical practise. It is possible that machine learning may detect nonlinear relationships in high-dimensional data. Machine learning refers to the process whereby computers may pick up new skills without any human input. In a number of health-related areas, state-of-the-art was bested by deep learning, a state-of-the-art machine learning approach [7]. Without the use of feature engineering, deep neural networks (DNNs) can accurately portray raw data [8].

There is a lack of research in the literature on the use of deep learning in the field of diabetes. Particularly for the treatment of diabetic eye disorders, deep learning has demonstrated to be a promising tool [9]. Therefore, the purpose of this study is to explore cutting-edge deep learning approaches to diabetes care.

II. DEEP LEARNING OVERVIEW

Healthcare and diabetes deep learning approaches are reviewed. Deep learning started with brain-like ANNs [10]. Fig. 2 shows an ANN's nodes and three layers: input, hidden, and output. These layers simulate neurones using arithmetic. Back-propagation can teach ANNs perceptions, but they can't generalise beyond supervised tasks. Hidden layers help DNNs generalise, collect data, and learn representations with hundreds of thousands of parameters. Over the previous two decades, computer technology and software have made DNN models more complicated [10]. Fig. 2 displays five diabetes research DNN topologies with nodes, cells, and connections. Deep learning libraries include Theano, Caffe, TensorFlow, CNTK, and PyTorch. Supporting many languages and hardware acceleration, these frameworks simplify DNN model creation.

Supervised, unsupervised, and reinforcement deep learning algorithms exist. Supervised learning classifies and regresses labelled input data during iterative model optimisation and backwards propagation. Diabetes research uses supervised learning-based DNNs such DMLPs, CNNs, and RNNs.

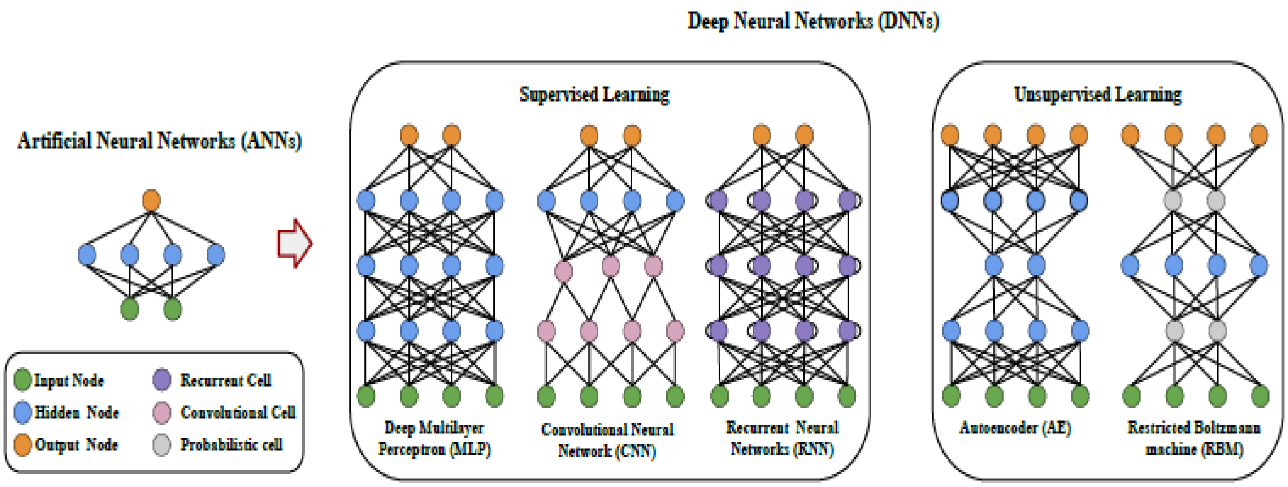

Figure 2: The visualization of ANNs and DNNs.

DMLP, a feed-forward-only (FC) neural network, is used in several DNN models. "Deep" modes have more than three layers since multilayer perceptrons are ANNs or DNNs. Nonlinear weight vectors, bias scalars, sigmoid, tanh, and ReLU activation functions [10] are all features of DMLPs. Multi-dimensional array data may be visualised by CNNs thanks to their convolutional layers. Most CNNs have a mapping capability for subsampling pools. Figure 2 shows how back-propagation helps train better convolutional models. In the ImageNet database and in industrial image identification applications, GPUs and TPUs have employed CNN-based models. AlexNet, VGGNet, Inception, and ResNet are all common CNN architectures.

Time-related data from successive signals is where RNNs really shine. During back-propagation, vanilla RNNs experience gradient vanishing and explosion [11]. Innovative RNN cells including long short-term memory (LSTM) and gated recurrent units (GRUs) [12] benefit from gate functions and long-term information persistence. Prediction and regression issues in natural language processing and speech identification were solved by RNN-based models. RNN models benefit from attention because it allows them to zero in on input sequences and map dependencies independently of physical distance.

Classless models are produced through unsupervised learning. Hidden structures and representations in input datasets are automatically revealed. Preprocessing, clustering, density estimation, and dimensionality reduction are all possible with unsupervised learning. Minimal Boltzmann machines and autoencoders. Adverse effects change. In probability theory, RBMs are used to create maps of probability distributions. Only training using bipartite connections between RBM neurones is more efficient. DBNs are the result of stacking RBMs [16]. Unsupervised learning may be spotted using DBNs. To improve performance on learning tasks, supervised learning may be used to fine-tune network weights [17]. When deep neural networks (DNNs) simulate policy, value-function, or system models, DRL outperforms humans. Networking instruction in a simulated environment [18].

III. METHODOLOGY

Deep learning for diabetes research was evaluated using PubMed, DBLP, and IEEEXplore. PubMed focuses on medical and biological studies, whereas DBLP is dedicated to computer science. Technical and scientific journals may be found in the IEEE Xplore database. These three databases do not need institutional subscriptions to use their search tools and user interfaces, unlike Ovid, Scopus, and Web of Science. We utilised freely available search engines to ensure the accuracy of our findings.

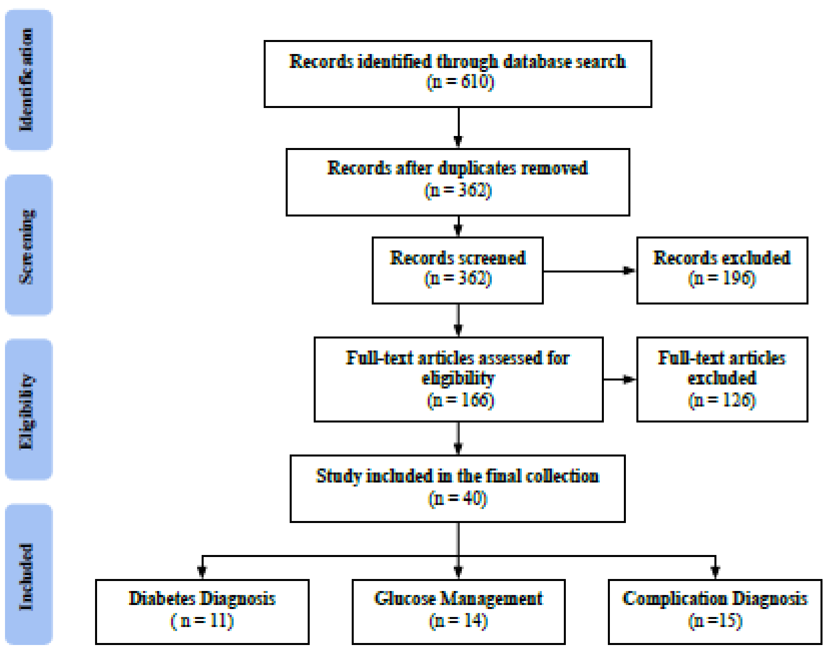

Figure 3: PRISMA flow of selection process.

A. Search Strategies

The keywords "diabetes," "glucose," and "artificial pancreas" were combined with the deep learning concepts using the Boolean operators AND and OR in our paper search. A specific search was conducted using the following terms: ((diabetes OR glucose OR artificial pancreas) AND (deep learning OR deep neural network OR convolution neural network OR convolutional neural network OR recurrent neural network OR LSTM OR autoencoder OR boltzmann machine OR deep belief network)). After amassing the results of a preliminary collection of relevant articles, we removed duplicates from multiple sources and then personally evaluated the remaining to evaluate them based on inclusion criteria.

B. Inclusion and Exclusion Criteria

This analysis synthesises primary, open-access research on the use of deep learning for diabetes management. The articles were organised according to three topics: diabetes diagnosis, glucose control, and complication diagnosis. Selected studies were able to conduct at least one of the following: give datasets and data processing details; provide detailed descriptions of technique; assess model performance using widely used metrics; and provide examples of the structure of DNNs.

The prevalence of diabetes-related retinopathy in published works is noteworthy. Therefore, we focused on research that either used large clinical data sets or revealed novel findings on DNNs. All posters, abstracts, methods reports, and reviews were rejected.

C. Information Extraction

We thoroughly examined the articles to get the information we need to assess deep learning systems. Each research was given a visual inspection before being included into one of the three tables below.

One Potential Scenario We initially synthesised the many application situations for each research in order to determine the distinct foci of each. To distinguish between studies involving persons with type 1 and type 2 diabetes, we've put y for type 1 diabetes research and z for type 2 diabetes research.

Examples, Part 2: In Part II, we examine the most popular model topologies and discuss the various DNN layer types. Additionally explored are ensemble approaches and hybrid designs.

The performance of deep learning algorithms depends on how accurate the data is. The generalizability of DNN models has been the subject of several studies utilising a variety of datasets (both public and private). The datasets utilised are thus briefly described in this section, together with information about their suppliers, categories, and file formats. To make it easier for future researchers interested in investigating these data scarcity issues to locate them, we have annotated publically available datasets with a question mark (?).

The fourth part, which is devoted to the development process, discusses the testing and validation stages as well as the initial construction of deep learning models. Deep learning is excellent at extracting representations from unstructured data, but the need for precise planning throughout the development stages compromises the models' utility and reproducibility.

Critical Findings No. 5: Below, we go through some key discoveries as well as the metrics and standards that will be used to judge how well we accomplished those goals. In comparison to sensitivity, specificity, and AUC, root mean square error (RMSE) is more often used in glucose control to diagnose diabetes and its effects. The results of the Cases are often congruent with the suggested explanations.

A Word About Baselines The bulk of these studies evaluated DNN algorithms using various benchmarking methodologies. Common statistical and machine learning methods include logistic regression (LR), autoregression (AR), autoregressive integrated moving average (ARIMA), supporting vector machines (SVM), random forests (RF), naive Bayes (NB), k-nearest neighbours (KNN), latent variable model (LVX), principal component analysis (PCA), and decision trees (DT). Additionally supplied for evaluation is the maximum baseline performance as measured by the Main Outcomes-aligned indicators.

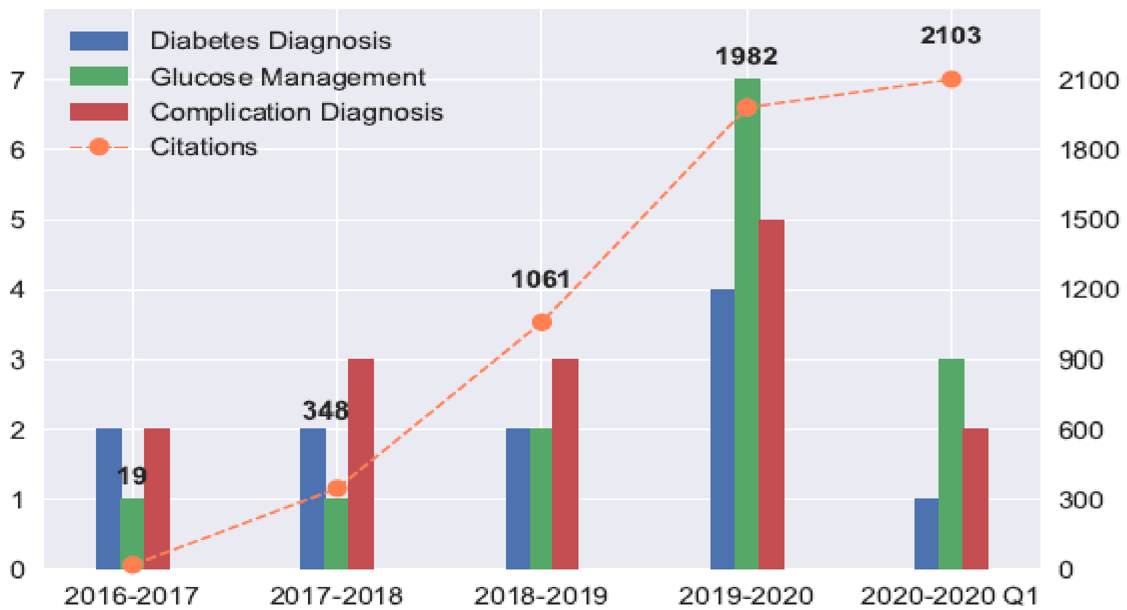

Figure 4: Number of articles included in the collection grouped by the year of publication and application filed.

As a review for a developing strategy like deep learning, Category 7 summarises the restrictions discovered in the selected research. These restrictions might motivate greater study aimed at enhancing the efficiency with which learning is applied across all fields.

IV. RESULTS

In Figure 3, we can see the sum of the articles found in the three databases used in the primary search: PubMed (307), DBLP (31) and IEEE Xplore (272). When all the duplicates were taken out, there were just 362.

We then categorised the papers using the standards we'd set. Forty articles made the cut after we carefully examined their complete texts to ensure they met our inclusion criteria. Our remaining dataset was partitioned into three groups based on the use cases: (n = 11) diabetes diagnosis, (n = 14) glucose control, and (n = 15) complications diagnosis. Figure 4 shows that majority of the articles under consideration for inclusion were published during the preceding two years, proving that deep learning research for diabetes is a relatively new field that is attracting a growing amount of attention. Google Scholar citation counts as of October 2020 were also computed and shown next to each of the featured works. Tables I, II, and III list the chosen works in reverse chronological order.

A. Diagnosis of Diabetes

Diabetes patients who get early diagnosis and treatment may profit considerably. Clinicians often utilise glucose-based haemoglobin A1c (HbA1c) testing to confirm diabetes [3]. Due to rural population and medical shortages, undiagnosed cases are common and predicted to grow [22]. T2D individuals without diabetic symptoms may develop chronic organ failure.

Population screening and noninvasive techniques are required to detect or predict diabetes. Table I shows the latest deep learning diabetes diagnostic decision-support algorithms. DMLP models are the most popular supervised and unsupervised learning approach applications. DMLP's feed-forward architecture and simple connections help binary classifiers on EHRs, while AEs and RBMs reveal data patterns unsupervised. Several studies have utilised UCI's Pima Indian Diabetes (PID) dataset [23]. 768 instances had eight traits and a binary diabetes classification. illustrates this dataset. Machine learning research uses the Pima Indians (PID) dataset because they have the highest T2D rate. This dataset with uniform measurements makes it easy to compare machine learning research. Table I includes Mount Sinai Data Warehouse and Practise Fusion dataset diabetes diagnostic EHR datasets. [23]

These datasets have an ICD-9 and ICD-10 diagnostic coding scheme despite their different sources. These technologies let researchers identify electronic health data of diabetics and their comorbidities. Deep learning on the PID dataset is limited by the lack of patients and characteristics. DNN generalisation must be proven using a large population dataset. Miott et al. used Deep Patient, a stack of denoising AEs, to learn representations from a big dataset. Area under the curve showed 0.907 diabetes classification accuracy [11]. A recent Ryu et al. study included 11,456 individuals [21]. A DMLP model screened for undiagnosed diabetes with an AUC of 80.11%. These studies extracted patient-specific traits using feature analysis and data normalisation.

Non-invasive diabetes detection is also promising. Lekha et al. [14] built a 1-D CNN architecture for real-time breath biomarker analysis for diabetes diagnosis and categorisation. MOS sensors collected breath samples for VOC measurement. The sensor array scanned a small gas chamber every 1000 seconds. After that, the signals were processed by the CNN classifier, which, unlike PCA, SVM, and SVD, may remove feature selection and maximise performance. Diabetes was diagnosed using ECG heart rate variability. Forty people responded in 10 minutes.

Table 1: Summary of selected articles from the literature on diabetes diagnosis.

Ref.	Cases	Models	Main Outcomes
[11]	Classification of diabetes	Denosing AE	AUC: 0.907
[12]	Prediction of diabetes	Modified LSTM, attention pooling layer	Precision of diagnosis, intervention, unplanned readmission: 66.2%, 78.7%, 79.0%
[13]	Detection of diabetesy	RBM and RNN	Sensitivity and precision: 90.66%, 75%

[14]	Prediction of diabetesy	Modified 1-D CNN and FC layers	AUC of T1D, T2D, healthy subjects: 0.9659, 0.9625, 0.9644
[15]	Detection of diabetes	5-layer CNN, LSTM, and SVM	Validation accuracy: 95.70%
[16]	Detection of diabetes	DMLP with dropout	Accuracy: 88.41%
[17]	Prediction of diabetes	DMLP	AUC without and with HbA1c: 0.703, 0.840
[18]	Prediction of the onset T2D	DMLP and a linear model	Sensitivity: 31.17%, AUC: 84.13%
[19]	Detection of diabetes	2 layer AE and a softmax layer	Sensitivity: 87.92%, specificity: 83.41%, accuracy: 86.26%
[20]	Prediction of diabetes	DBN	Sensitivity: 100%, F1 score: 0.808
[21]	Detection of undiagnosed diabetes	2 hidden layer DMLP with dropout	AUC: 80.11%

The ECG data were gathered at a frequency of 500 Hz and then processed using digital bandpass filtering and thresholding for real-time detection. Researchers employed the Pan-Tompkins method to get the heart rate time, and then combined CNN, LSTM, and SVM into a single hybrid deep learning model. They achieved a validation accuracy of 95.7% using this model.

B. Glucose Management

Managing glucose in diabetes means avoiding low and high blood sugar, or hypo- and hyperglycemia, respectively. As can be seen in Figure 1, the rapid progress of deep learning has been greatly aided by the digitisation of diabetic self-management. Blood glucose prediction, abnormal blood glucose detection, insulin administration control, and decision assistance are all subfields within the larger topic of glucose management.

Predicting blood glucose levels is a hot issue right now. By doling out the right amounts of insulin and/or glucagon based on an accurate BG forecast, sensor-enhanced insulin pumps (like predictive low-glucose insulin suspension) and AP systems (like model predictive control) may help reduce the likelihood of BG abnormalities (like hypoglycemia and hyperglycemia). Smartphone apps may help people with diabetes monitor the factors in their surroundings that impact their blood sugar levels. A multimodal time series may be constructed and examined by deep learning algorithms by linking CGM measurements with other self-reported events, such as meal composition and insulin administration. In general, forecasts with a PH of 30 minutes are considered to be short-term, while those with a PH of 60 minutes or more are considered to be long-term.

The RNN-based architecture is useful here because of its shown success in temporal sequence processing and regression. Table II demonstrates that RNNs supplemented with LSTM cells are superior at predicting glucose levels. For better 30- and 60-minute prediction than the engineering physiological model (EPM), Mirshekarian et al. [25] used support vector regression (SVR). The EPM is a continuous dynamic model that identifies system states by considering digestion dynamics, insulin digestion dynamics, and glucose-insulin dynamics. Memory-based case-based prediction was modelled using a neural attention layer [32].

TABLE II: Summary of selected articles from the literature on glucose management.

Ref.	Cases	Models	Main Outcomes

[24]	Detection of hypoglycemia	DBN by stacking RBMs	Sensitivity: 79.70%, specificity: 50.00%
[25]	Prediction of BG levels	LSTM and a linear layer	RMSE for 30, 60-min PH: 21.4, 38.0 mg/dL
[26]	Prediction of BG levels	LSTM, Bidirectional LSTM and 3 FC layers	RMSE for 30, 45, 60-min PH: 11.63, 21.75, 36.92 mg/dL
[27]	Prediction of BG levels	Deep sequential polynomial multi-output model (RNN)	Absolute percentage error for 30-min PH: 4.87
[28]	Glycemic control	CNN (Inception-v3)	Time in rage (TIR) of 70-180 mg/dL: 91.76%, top-1 accuracy of the image classification: 81.65%
[29]	Prediction of BG levels	LSTM with dynamic time warping	Clark Error Grid zones of next-day PH (A: 84.12, B: 15.16, C: 0, D: 0.72, E: 0)%
[30]	Prediction of BG levels	CNN, LSTM and 2 FC layers	RMSE for 30, 60-min PH: 9.38, 18.87 (1); 21.07, 33.27 (2) mg/dL
[31]	Prediction of BG levels	LSTM and a FC layer	RMSE for 30, 45, 60-min PH: 19.47, 26.47, 32.38 mg/dL
[32]	Prediction of BG levels	Memory- Augmented LSTM with neural attention weights	RMSE for 30, 60-min PH: 18.74, 30.63 (1); 1.23, 2.27 (2); 2.93, 4.92 (3) with input of CGM, insulin and meal events
[33]	Prediction of BG levels	Dilated CNN (residual and parameterized skip connections)	RMSE for 30, 60-min PH: 8.88, 19.90 (1); 19.19, 31.78 (2); 19.28, 31.83 (3)
[34]	Glycemic control	Deep Q-network with GRU or 1-D CNN	Average risk index for the virtual subject: 9.26
[35]	Prediction of BG levels	2 branches of LSTM cells (past and future information)	Average RMSE for PH of 60 minutes: 11.72 (1), 21.09 (2)
[36]	Prediction of HbA1c	1-D CNN, Inception module, FC layers	Mean absolute error: 4.80, the coefficient of determination: 0.71
[37]	Prediction of BG levels	LSTM and 2 FC layers	RMSE for 30, 60-min PH: 18.867, 31.403

CNNs assess meal macronutrients for human nutrition [28]. The Food-101 dataset of food photos may be used to train a CNN model for decision support and AP systems to calculate meal bolus insulin. The UVA/Padova type 1 diabetes simulator was used to assess the algorithm's dependability under realistic conditions including meal size and carbohydrate content. The FDA has authorised the UVA/Padova T1D simulator, a glucose-insulin dynamics simulator created by the University of Virginia (US) and Padova (Italy). Due to the high costs and safety issues of clinical trials on people and animals, many research organisations have turned to computer simulation, or in silico, to test algorithms in various virtual conditions.

Fox et al. investigated GRU and 1-D CNN DRL algorithms for basal insulin management using the UVA/Padova T1D simulator [27]. Recent studies have studied modern DRL algorithms for glucagon and bolus insulin, two important hormones involved in glycemic control [87, 88]. Glucose forecasting relies on the simulator to create population data sets for preliminary verification. The 2018 BG level prediction challenge used the OhioT1DM dataset, which was updated for 2020 [38]. Despite most research tests on their own proprietary clinical databases. Twelve type 1 diabetics provided eight weeks of multi-modal data (continuous glucose monitor, diet, insulin, and physical activity). An unsupervised learning system using DBNs and ECG data identified hypoglycemia in T1D children [24]. CNN-LSTM models were used to identify sleep-related hypoglycemia in electrocardiograms [39].

C. Diagnosis of Complications

Table III shows that most research has concentrated on medical imaging's potential for identifying diabetes-related problems at an early stage. Monitoring diabetes is a time-consuming ordeal that requires regular, individualised visits to the doctor's office [55]. Diabetes therapy increases healthcare expenses and wait times. For widespread monitoring and surveillance, communities need automated methods of screening for, detecting, and diagnosing issues related to diabetes. Diabetic retinopathy (DR) is the most common cause of blindness in people with diabetes. Diagnosis of DR before to blindness is very uncommon.

The good news is that deep learning algorithms can compete with humans in DR problem solving [9]. After the success of the CV, several more models were developed utilising convolutional neural networks (CNNs) to characterise retinal fundus pictures. In 2015, the Kaggle DR screening competition was won by CNNs trained on a publicly accessible dataset. Databases like Messidor-2 and E-Ophtha have DR inspection images. Table III demonstrates that, throughout 11 tests, 91% of the time, CNNs successfully detected DR. ElTanboly et al. [42] used a multistage deep fusion classification network with a stack of non-negativity-constrained AEs to identify DR in retinal OCT images that seemed normal at first glance. The AE model performed well when applied to a dataset consisting of 52 people. The two most popular CNN research architectures are VGGNet (used by 4% of researchers) and Inception (used by 5%).

VGGNet-s performed best on the Kaggle dataset in [47]. While VGGNet, developed at the University of Oxford (UK), enhances ImageNet recognition performance using a moderate kernel size and deep networks, Inception uses sparse connections between activation functions in an Inception module to maximise GPU computing efficiency.

Both approaches have a high degree of success in detecting DR. With a 96% sensitivity for referable DR, Abramoff et al. [40] used a VGGNet-based model to identify many types of DR on Messidor-2. Two large datasets [44, 49] with people of varying racial and ethnic backgrounds verified the VGG-adapted structure. They found that deep learning systems were able to detect the DR with more accuracy and speed than humans. Gulshan et al.'s Inception-based architecture accurately identified cases of referable DR with a sensitivity of 96.8 and a specificity of 87.0. These systems have been studied in outpatient settings using a variety of assessment criteria. The Inception-like strategy tried and true by Ruamviboonsuk et al. [48] is now being used on a countrywide scale. When compared to human professionals, deep learning is more insightful, but it can't match their level of accuracy. Their efforts have been acknowledged as comparable to those of humans in the 2019 publication of the Artificial Intelligence Index [56].

TABLE III: Summary of selected articles from the literature on diagnosis of complications.

Ref.	Cases	Models	Main Outcomes
[40]	Referable DR detection	CNN (Inspired by AlexNet, VGGNet)	Sensitivity: 96.8%, specificity: 87.0%, AUC: 0.980
[41]	Referable DR detection	CNN (Inception-v3), an ensemble of 10 networks	Sensitivity: 97.5% (1), 96.1% (2), specificity: 93.9% (1), 93.4% (2), AUC: 0.990 (1), 0.991 (2)
[42]	DR detection	A stack of non-negativity-constrained AEs	Sensitivity: 92%, specificity: 83%, accuracy: 100%
[43]	DR detection	Customized CNN: (5 residual blocks), DT classifier	Sensitivity: 94% (1), 93% (2), 90% (3), specificity: 98% (1), 87% (2), 94% (3), AUC: 0.97 (1), 0.94 (2), 0.95 (3)
[44]	Referable DR detection	CNN (Adapted VGGNet)	Sensitivity: 90.5%, specificity: 91.6%, AUC: 0.936
[44]	Referable DR detection	CNN (Inception-v3)	Sensitivity: 92.3%, specificity: 93.7%, 96% of participants satisfied with the model
[46]	Moderate or worse DR detection	CNN (Inception-v4), an ensemble of 10 networks	Sensitivity: 97.1%, specificity: 92.3%, AUC: 0.986
[47]	DR detection	CNN (VGGNet-s)	Sensitivity: 86.47%, specificity: 97.43%, AUC: 0.9786, accuracy: 95.68%
[48]	Referable DR detection	CNN (Inception-v4), an ensemble of 10 networks	Sensitivity: 96.8%, specificity: 95.6%, AUC: 0.987
[49]	Referable DR detection	CNN(Adapted VGGNet)	The estimation of DR prevalence: 16.1%, the AUC for referable DR: 0.863, the time taken to diagnose: 10.4h, risk factor: 0.743
[50]	Estimation of DR severity scale	CNN pillars (Inception-v3) and RF	AUC at month 6, 12, 24: 0.68, 0.79, 0.77
[51]	Prediction of mortality in ICU	1-D CNN and 2 FC layers	AUC: 0.885
[52]	Prediction of myocardial infarction	DMLP	AUC: 0.767, with hazard ratio: 0.81 and 0.63
[53]	Classification of diabetic foot	Customized 9-layer CNN	Sensitivity: 0.9167, AUC: 0.8533
[54]	Detection of diabetic neuropathy	U-Net CNNs (5 ensembles)	Fibre length 0.933, length/segment: 0.656, branch points: 0.891, nail points: 0.623

The research also highlights deep learning's flexibility in addressing different problems. With an AUC of 0.885, Wittler et al. [51] developed a CNN-based model to predict mortality from ICU patient data. The MIMIC-III ICU dataset used in this analysis is publicly available and may be downloaded for free [57]. Using a U-Net convolutional neural network (CNN) trained on a publicly available dataset was recommended by Williams et al. [54] to aid in the identification of diabetic neuropathy. The researchers suggest that this approach might be used in therapeutic contexts, and their findings demonstrate excellent localisation performance for quantitative assessment. Using a publically available dataset, a CNN was developed in [53] to detect plantar ulcers in thermographic images of diabetic feet.

In addition, Yamada et al. [52] used a DMLP model that outperformed standard LR analysis to evaluate the risk of cardiovascular disease in order to do a comparison of three anti-diabetic medicines. Finally, it's worth noting that almost all studies looked at microvascular outcomes including DR [40– 50], diabetic foot [53], and diabetic neuropathy [54], whereas just one looked at macrovascular concerns such cardiovascular illnesses [52].

D. Summary of Deep Learning Techniques

The field of diabetes research has taken to using both supervised and unsupervised deep learning architectures.

Clinical imaging challenges are a common application of convolutional neural network-based systems. When it comes to feature extraction from raw data, CNNs shine [8]. Previously, this would have needed either specialised knowledge in the field of image processing or a custom-built solution. Medical image and scan interpretation for the purpose of detecting diabetes-related problems is where CNNs have so far seen the most widespread use. Another potential use of CNN-based systems that might be useful to patients with diabetes is the prediction of macronutrients from food photos. The CV community is always looking at new methods to boost model efficiency while simplifying the system. With the advent of more sophisticated CNN configurations (VGGNet, Inception), more robust methods have been developed. In addition, 1-dimensional convolutional neural networks (CNNs) have been investigated for their potential to analyse sequential signals by using convolutional filters with a wide receptive field to extract data features. The convolutional RNN (CRNN) [30] and the CNN-LSTM [15], [39] are two examples of hybrid models that take use of the superior sequence processing skills of RNNs by using LSTM layers to interpret the input and calculate the temporal connections. In particular, RNN-based designs are currently considered state-of-the-art in BG prediction within the context of diabetes management apps. A real-time glucose map may be generated from CGM data using modern recursive algorithms and intricate cell topologies. Recent developments in natural language processing (NLP) have been the subject of several academic investigations. Neuronal attention mechanisms [21, 32] and bidirectional LSTM [26] are two such examples.

Diabetic diagnosis often employs unsupervised learning algorithms like DMLP. Due to the variety of record formats included in EHR datasets, however, performing such tasks generally requires meticulous feature selection and normalisation in pre-processing. The most important characteristics are isolated using standard machine learning methods.

For each data characteristic, scores were calculated using principal components analysis with weights and coefficients as shown in [20]. By evaluating correlations between non-invasive measurements and person traits, LR analysis was used to identify key factors in diabetes diagnosis [21].

Due to its adaptability and modular structure, DNN layers may be included into a variety of different models. In addition to recurrent neural networks [50], deep turings [43], and recurrent neural networks [50], the linear model [18], support vector machines [15], and deep turings [18] all incorporate data features at the input or perform a second-level analysis at the output, making them hybrid learning models. References [41, 46, 48, 50] and [54] further detail deep learning ensemble models. The ensemble is used to generate test results by linearly averaging the outputs of many individually trained CNNs on the same dataset. In order for each CNN to learn its own representation and enhance accuracy and generalisation, it is seeded with a different set of data at the beginning of training.

Since millions of parameters in the DNN units must be fine-tuned during training, doing so from start is a time-consuming process. A time-saving technique that might be utilised to deal with this is pre-training, also known as transfer learning. There is evidence in [28, 41] that the ImageNet database is used as a supplementary portion of CNN pre-training, especially for medical imaging applications. Convergence on target datasets may be accelerated by fine-tuning the weights in accordance with the ImageNet principles, provided adequate computing resources are available [58]. When it comes to pretraining for glycemic control tasks [25, 26, 32], we employ in silico datasets obtained from simulators in addition to some real clinical data. Efficient in accommodating the massive data needs of DNN training. To perform the discriminative fine-tuning described in [20], DBNs first conduct an unsupervised pre-training phase to determine suitable starting weights. Data augmentation is described in depth in [47, 53] as an alternate strategy for enhancing model performance with sparse data. These experiments rotated, flipped, and rearranged the photos they used to create bigger training datasets.

V. DISCUSSION

A. Limitations and Challenges

While deep learning has made significant strides in many diabetes-related fields, it still lacks definitive proof of safety and efficacy for widespread use in healthcare systems. Prior to its widespread implementation in actual therapeutic settings, deep learning still faces a number of hurdles.

Table IV provides a summary of the restrictions placed on the data in terms of quantity, variation, quality, feature processing, and interpretability. Real-world data from diabetic patients is likely to be erroneous because to human error and sensor distortions.

It might be expensive and time-consuming to gather reliable information. It's possible that data privacy restrictions may hinder collaboration between academic institutions when sharing information. As a result, many research suffer from a too-small sample size. The complexity of glucose dynamics makes it hard to analyse the data available to characterise people with diabetes.

TABLE IV: Summary of the limitations and challenges identified by the selected articles.

Category	Description	References
Data Volume	Training a deep model for complicated tasks requires a high volume of data. Collecting data from people with diabetes is often time-consuming and expensive, compared to other tasks in CV and NLP. Consequently, many studies face a shortage of data during their research cycles.	[12], [17], [29], [37], [40], [50], [53]
Data Variability	The variability among people with diabetes is large due to the complex glucose dynamics. To obtain better generalization for deep learning models, the training data needs to cover a diverse range of individuals, such as people of different ages and comorbidities. However, many	[11], [21], [24], [28], [36], [46], [48], [50], [52]

	datasets are often collected from a specific cohort of people, which lacks diversity and could bring bias to the learning.	
Data Quality	Similar to many other problems in healthcare, most of the diabetes datasets are heterogeneous, sparse, and noisy with some missing values. It is not realistic to collect perfect data from either clinical practice or daily self-management, e.g. the unavoidable errors from CGM sensors.	[11], [29], [36], [37], [40], [51]
Feature Processing	The major challenge in feature processing is to find the most effective features for models to learn the representations. Manually screening and analyzing each feature in a diabetes dataset could require a lot of engineering work, but using automated data-driven methods, such as PCA, would ignore some physiological knowledge and rely too much on the characteristics of the data. A more comprehensive analysis of additional factors and features is needed with the advances in data collecting and physiological models.	[18], [25], [26], [41], [49], [52]
Interpretability	The interpretability, i.e. explainability, stands for how the model obtains the corresponding output based on a set of inputs. It is an important goal for AI applications in healthcare to convince clinicians to adopt such systems. In many cases, deep learning models are regarded as" black boxes" with a lack of model transparency due to complex nonlinear layers. As a consequence, if the model performance degrades in certain circumstances, it might be difficult to explain why.	[30], [40], [44], [48], [50]

In addition, deep learning algorithms tend to be secretive. When making potentially fatal judgements, doctors must have full confidence in the accuracy of the models they're using. In order to learn efficient patterns from non-linear data, DNN layers have complicated structures, however this comes at the expense of the model's interpretability. Therefore, it is important to assess the trade-off between performance and interpretability when investigating deep learning for diabetes.

The efficiency with which deep learning models may be trained is expected to improve as a result of new algorithmic and hardware developments [13, 53].

B. Opportunities and Future Work

To make sense of the medical applications of deep learning technologies, the AI community has recently focused on increasing model openness and gaining an understanding of model working. In particular, a unified framework, the SHapley Additive exPlanations (SHAP), has been proven by various data-driven applications in the healthcare sector to characterise the input features that contribute to the final output [59].

This method may also be used to choose input qualities by ranking them in order of importance. The outcomes of a CNN research credited with SHAP analysis are shown in Table III [50]. To better comprehend the learned properties of CNN layers, t-distributed stochastic neighbour embedding (t-SNE) was used to show the clusters of heartbeat data according to glucose levels in [39]. Qualitative study of the produced feature maps using t-SNE may also be useful for other CNN applications, such as DR detection. The dynamics of glucose and insulin have recently been studied, and the results have been shown to be congruent with neural network models [60]. DNN performance may be evaluated and interpretability can be enhanced using similar approaches.

If data-driven models are reinforced with expert knowledge during the learning process, it may be possible to get a deeper understanding of the underlying mechanisms of a health issue like diabetes. In specifically, two strategies might be feasible options.

One is to factor in previously acquired knowledge as a training aid, and the other is to use physiological parameters as a model input. Expertise is also necessary for the creation of safety restrictions and the computation of confidence in model outcomes.

The necessity for further confirmation of their findings in real-world conditions was a major factor in the selection of many publications [35, 36, 44, 50, 54]. One Google team has made some headway in this area. Research focused on patients with diabetic eye problems was conducted in 11 clinics using deep learning [61]. The research indicates that a number of social and environmental conditions must be satisfied before such automated systems may be extensively used.

VI. CONCLUSION

In this research, we examine how scientists are actively using deep learning methods to further their investigation of diabetes. We conducted a comprehensive literature search, selected a sample of relevant papers, and synthesised the most significant data across three domains: diabetes diagnosis, glucose management, and repercussions. Experiments conducted on these challenges utilising various DNN architectures and learning strategies have shown better results than those possible using conventional machine learning techniques. However, many challenges have been identified in the literature. These include data accessibility, feature processing, and model interpretability. Applying state-of-the-art deep learning technology to massive, multi-modal data sets linked to diabetes treatment shows tremendous potential for addressing these difficulties in the future. We believe that deep learning technologies will soon become widespread in clinical settings, vastly improving diabetes treatment overall.

REFERENCES

[1] P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova et al., "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas," Diabetes research and clinical practice, vol. 157, p. 107843, 2019.

[2] Global report on diabetes. World Health Organization, 2016.

[3] American Diabetes Association and others, “2. classification and diagnosis of diabetes,” Diabetes care, vol. 40, no. Supplement 1, pp. S11–S24, 2017.

[4] D. Rodbard, “Continuous glucose monitoring: a review of successes, challenges, and opportunities,” Diabetes technology & therapeutics, vol. 18, no. S2, pp. S2–3, 2016.

[5] A. Slomski, “Artificial pancreas improves glycemic control,” Jama, vol. 320, no. 20, pp. 2068–2068, 2018.

[6] F. L. Schwartz, C. R. Marling, and R. C. Bunescu, “The promise and perils of wearable physiological sensors for diabetes management,” Journal of diabetes science and technology, vol. 12, no. 3, pp. 587–591, 2018.

[7] D. Rav`ı, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, “Deep learning for health informatics,” IEEE journal of biomedical and health informatics, vol. 21, no. 1, pp. 4–21, 2016.

[8] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[9] A. Grzybowski, P. Brona, G. Lim, P. Ruamviboonsuk, G. S. Tan, M. Abramoff, and D. S. Ting, “Artificial intelligence for diabetic retinopathy screening: a review,” Eye, pp. 1–10, 2019.

[10] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.

[11] 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,” Scientific reports, vol. 6, no. 1, pp. 1–10, 2016.

[12] T. Pham, T. Tran, D. Phung, and S. Venkatesh, “Predicting healthcare trajectories from medical records: A deep learning approach,” Journal of biomedical informatics, vol. 69, pp. 218–229, 2017.

[13] S. Ramesh, H. Balaji, N. C. S. Iyengar, and R. D. Caytiles, “Optimal predictive analytics of Pima diabetics using deep learning,” International Journal of Database Theory and Application, vol. 10, no. 9, pp. 47–62, 2017.

[14] S. Lekha and M. Suchetha, “Real-time non-invasive detection and classification of diabetes using modified convolution neural network,” IEEE journal of biomedical and health informatics, vol. 22, no. 5, pp. 1630–1636, 2017.

[15] G. Swapna, R. Vinayakumar, and K. Soman, “Diabetes detection using deep learning algorithms,” ICT Express, vol. 4, no. 4, pp. 243–246, 2018.

[16] A. Ashiquzzaman, A. K. Tushar, M. R. Islam, D. Shon, K. Im, J.-H. Park, D.-S. Lim, and J. Kim, “Reduction of overfitting in diabetes prediction using deep learning neural network,” in IT Convergence and Security 2017. Springer, 2018, pp. 35–43.

[17] S. Sp¨anig, A. Emberger-Klein, J.-P. Sowa, A. Canbay, K. Menrad, and D. Heider, “The virtual doctor: An interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes,” Artificial intelligence in medicine, vol. 100, p. 101706, 2019.

[18] B. P. Nguyen, H. N. Pham, H. Tran, N. Nghiem, Q. H. Nguyen, T. T. Do, C. T. Tran, and C. R. Simpson, “Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records,” Computer methods and programs in biomedicine, vol. 182, p. 105055, 2019.

[19] K. Kannadasan, D. R. Edla, and V. Kuppili, “Type 2 diabetes data classification using stacked autoencoders in deep neural networks,” Clinical Epidemiology and Global Health, vol. 7, no. 4, pp. 530–535, 2019.

[20] P. Prabhu and S. Selvabharathi, “Deep belief neural network model for prediction of diabetes mellitus,” in 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC). IEEE, 2019, pp. 138–142.

[21] K. S. Ryu, S. W. Lee, E. Batbaatar, J. W. Lee, K. S. Choi, and H. S. Cha, “A deep learning model for estimation of patients with undiagnosed diabetes,” Applied Sciences, vol. 10, no. 1, p. 421, 2020.

[22] N. Cho, J. Shaw, S. Karuranga, Y. Huang, J. da Rocha Fernandes, A. Ohlrogge, and B. Malanda, “IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045,” Diabetes research and clinical practice, vol. 138, pp. 271–281, 2018.

[23] D. Dua and C. Graff, “UCI machine learning repository,” 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[24] P. P. San, S. H. Ling, and H. T. Nguyen, “Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 3503–3506.

[25] S. Mirshekarian, R. Bunescu, C. Marling, and F. Schwartz, “Using LSTMs to learn physiological models of blood glucose behavior,” in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 2887–2891.

[26] Q. Sun, M. V. Jankovic, L. Bally, and S. G. Mougiakakou, “Predicting blood glucose with an LSTM and Bi-LSTM based deep neural network,” in 2018 14th Symposium on Neural Networks and Applications (NEUREL). IEEE, 2018, pp. 1–5.

[27] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, “Deep multi-output forecasting: Learning to accurately predict blood glucose trajectories,” in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 1387– 1395.

[28] A. Chakrabarty, F. J. Doyle, and E. Dassau, “Deep learning assisted macronutrient estimation for feedforward-feedback control in artificial pancreas systems,” in 2018 Annual American Control Conference (ACC). IEEE, 2018, pp. 3564–3570.

[29] S. H. A. Faruqui, Y. Du, R. Meka, A. Alaeddini, C. Li, S. Shirinkam, and J. Wang, “Development of a deep learning model for dynamic forecasting of blood glucose level for type 2 diabetes mellitus: Secondary analysis of a randomized controlled trial,” JMIR mHealth and uHealth, vol. 7, no. 11, p. e14452, 2019.

[30] K. Li, J. Daniels, C. Liu, P. Herrero-Vinas, and P. Georgiou, “Convolutional recurrent neural networks for glucose prediction,” IEEE journal of biomedical and health informatics, 2019.

[31] A. Aliberti, I. Pupillo, S. Terna, E. Macii, S. Di Cataldo, E. Patti, and A. Acquaviva, “A multi-patient data-driven approach to blood glucose prediction,” IEEE Access, vol. 7, pp. 69 311–69 325, 2019.

[32] S. Mirshekarian, H. Shen, R. Bunescu, and C. Marling, “LSTMs and neural attention models for blood glucose prediction: Comparative experiments on real and synthetic data,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2019, pp. 706–712.

[33] K. Li, C. Liu, T. Zhu, P. Herrero, and P. Georgiou, “Glunet: A deep learning framework for accurate glucose forecasting,” IEEE journal of biomedical and health informatics, 2019.

[34] I. Fox and J. Wiens, “Reinforcement learning for blood glucose control: Challenges and opportunities,” in Reinforcement Learning for Real Life (RL4RealLife) Workshop in 36th International Conference on Machine Learning (ICML), 2019.

[35] E. M. Aiello, G. Lisanti, L. Magni, M. Musci, and C. Toffanin, “Therapy-driven deep glucose forecasting,” Engineering Applications of Artificial Intelligence, vol. 87, p. 103255, 2020.

[36] A. Zaitcev, M. R. Eissa, Z. Hui, T. Good, J. Elliott, and M. Benaissa, “A deep neural network application for improved prediction of HbA1c in type 1 diabetes,” IEEE Journal of Biomedical and Health Informatics, 2020.

[37] J. Martinsson, A. Schliep, B. Eliasson, and O. Mogren, “Blood glucose prediction with variance estimation using recurrent neural networks,” Journal of Healthcare Informatics Research, vol. 4, no. 1, pp. 1–18, 2020.

[38] C. Marling and R. Bunescu, "The OhioT1DM dataset for blood glucose level prediction: Update 2020," in The 5th KDH workshop, ECAI 2020, 2020, pp. 71–74.

[39] M. Porumb, S. Stranges, A. Pescap`e, and L. Pecchia, "Precision medicine and artificial intelligence: A pilot study on deep learning for hypoglycemic events detection based on ECG," Scientific Reports, vol. 10, no. 1, pp. 1–16, 2020.

[40] M. D. Abr`amoff, 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," Investigative ophthalmology & visual science, vol. 57, no. 13, pp. 5200–5206, 2016.

[41] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," Jama, vol. 316, no. 22, pp. 2402–2410, 2016.

[42] A. ElTanboly, M. Ismail, A. Shalaby, A. Switala, A. El-Baz, S. Schaal, G. Gimel' farb, and M. El-Azab, "A computer-aided diagnostic system for detecting diabetic retinopathy in optical coherence tomography images," Medical physics, vol. 44, no. 3, pp. 914–923, 2017.

[43] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," Ophthalmology, vol. 124, no. 7, pp. 962–969, 2017.

[44] D. S. W. Ting, C. Y.-L. Cheung, G. Lim, G. S. W. Tan, N. D. Quang, A. Gan, H. Hamzah, R. Garcia-Franco, I. Y. San Yeo, S. Y. Lee et al., "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," Jama, vol. 318, no. 22, pp. 2211–2223, 2017.

[44] S. Keel, P. Y. Lee, J. Scheetz, Z. Li, M. A. Kotowicz, R. J. MacIsaac, and M. He, "Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study," Scientific reports, vol. 8, no. 1, pp. 1–6, 2018.

[46] J. Krause, V. Gulshan, E. Rahimy, P. Karth, K. Widner, G. S. Corrado, L. Peng, and D. R. Webster, "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy," Ophthalmology, vol. 125, no. 8, pp. 1264–1272, 2018.

[47] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," Computers & Electrical Engineering, vol. 72, pp. 274–282, 2018.

[48] P. Ruamviboonsuk, J. Krause, P. Chotcomwongse, R. Sayres, R. Raman, K. Widner, B. J. Campana, S. Phene, K. Hemarat, M. Tadarati et al., "Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program," NPJ digital medicine, vol. 2, no. 1, pp. 1–9, 2019.

[49] D. S. Ting, C. Y. Cheung, Q. Nguyen, C. Sabanayagam, G. Lim, Z. W. Lim, G. S. Tan, Y. Q. Soh, L. Schmetterer, Y. X. Wang et al., "Deep learning in estimating prevalence and systemic risk factors for diabetic retinopathy: A multi-ethnic study," NPJ digital medicine, vol. 2, no. 1, pp. 1–8, 2019.

[50] F. Arcadu, F. Benmansour, A. Maunz, J. Willis, Z. Haskova, and M. Prunotto, "Deep learning algorithm predicts diabetic retinopathy progression in individual patients," NPJ digital medicine, vol. 2, no. 1, pp. 1–9, 2019.

[51] I. Wittler, X. Liu, and A. Dong, "Deep learning enabled predicting modeling of mortality of diabetes mellitus patients," in Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning), 2019, pp. 1–6.

[52] T. Yamada, K. Iwasaki, S. Maedera, K. Ito, T. Takeshima, H. Noma, and N. Shojima, "Myocardial infarction in type 2 diabetes using sodiumglucose cotransporter-2 inhibitors, dipeptidyl peptidase-4 inhibitors, or glucagon-like

peptide-1 receptor agonists: Proportional hazards analysis by deep neural network-based machine learning," Current Medical Research and Opinion, no. just-accepted, pp. 1–1, 2019.

[53] I. Cruz-Vega, D. Hernandez-Contreras, H. Peregrina-Barreto, J. d. J. Rangel-Magdaleno, and J. M. Ramirez-Cortes, "Deep learning classification for diabetic foot thermograms," Sensors, vol. 20, no. 6, p. 1762, 2020.

[54] B. M. Williams, D. Borroni, R. Liu, Y. Zhao, J. Zhang, J. Lim, B. Ma, V. Romano, H. Qi, M. Ferdousi et al., "An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: A development and validation study," Diabetologia, vol. 63, no. 2, pp. 419–430, 2020.

[55] E. W. Gregg, N. Sattar, and M. K. Ali, "The changing face of diabetes complications," The lancet Diabetes & endocrinology, vol. 4, no. 6, pp. 537–547, 2016.

[56] R. Perrault, Y. Shoham, E. Brynjolfsson, J. Clark, J. Etchemendy, B. Grosz, T. Lyons, J. Manyika, and S. Niebles, "The AI Index 2019 annual report." 2019.

[57] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, a freely accessible critical care database," Scientific data, vol. 3, p. 160035, 2016.

[58] K. He, R. Girshick, and P. Doll´ar, "Rethinking ImageNet pre-training," in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 4918–4927.

[59] S. M. Lundberg, B. Nair, M. S. Vavilala, M. Horibe, M. J. Eisses, T. Adams, D. E. Liston, D. K.-W. Low, S.-F. Newman, J. Kim et al., "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," Nature biomedical engineering, vol. 2, no. 10, pp. 749–760, 2018.

[60] T. Kushner, S. Sankaranarayanan, and M. Breton, "Conformance verification for neural network models of glucose-insulin dynamics," in Proceedings of the 23rd International Conference on Hybrid Systems: Computation and Control, 2020, pp. 1–12.

[61] E. Beede, E. Baylor, F. Hersch, A. Iurchenko, L. Wilcox, P. Ruamviboonsuk, and L. M. Vardoulakis, "A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–12.