



# The Prediction Process Based on Deep Recurrent Neural Networks: A Review

Diyar Qader Zeebaree<sup>1</sup>, Adnan Mohsin Abdulazeez<sup>1</sup>, Lozan M. Abdullrhman<sup>2</sup>,  
Dathar Abas Hasan<sup>3\*</sup> and Omar Sedqi Kareem<sup>3</sup>

<sup>1</sup>Research Center, Duhok Polytechnic University, Kurdistan Region, Iraq.

<sup>2</sup>Duhok Polytechnic University, Kurdistan Region, Iraq.

<sup>3</sup>Shekhan Technical Institute, Duhok Polytechnic University, Kurdistan Region, Iraq.

## Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

## Article Information

DOI: 10.9734/AJRCOS/2021/v11i230259

### Editor(s):

(1) Dr. Hasibun Naher, BRAC University, Bangladesh.

### Reviewers:

(1) Ripon Patgiri, National Institute of Technology, India.

(2) Abdul Ghafar Bin Jaafar, Universiti Teknologi Malaysia, Malaysia.

Complete Peer review History: <https://www.sdiarticle4.com/review-history/72108>

Review Article

Received 02 June 2021  
Accepted 07 August 2021  
Published 20 August 2021

## ABSTRACT

Prediction is vital in our daily lives, as it is used in various ways, such as learning, adapting, predicting, and classifying. The prediction of parameters capacity of RNNs is very high; it provides more accurate results than the conventional statistical methods for prediction. The impact of a hierarchy of recurrent neural networks on Predicting process is studied in this paper. A recurrent network takes the hidden state of the previous layer as input and generates as output the hidden state of the current layer. Some of deep Learning algorithms can be utilized in as prediction tools in video analysis, musical information retrieval and time series applications. Recurrent networks may process examples simultaneously, maintaining a state or memory that recreates an arbitrarily long background window. Long Short-Term Memory (LSTM) and Bidirectional RNN (BRNN) are examples of recurrent networks. This paper aims to give a comprehensive assessment of predictions based on RNN. Additionally, each paper presents all relevant facts, such as dataset, method, architecture, and the accuracy of the predictions they deliver.

**Keywords:** Prediction; deep learning; RNN; LSTM; bidirectional recurrent neural network.

\*Corresponding author: E-mail: dathar.hasan@dpu.edu.krd;

## 1. INTRODUCTION

Accurate prediction is crucial for producing optimal results with deep RNN models over machine learning algorithms. Prediction is an essential factor in RNN Picture captioning, synthesizing expression, music generation, and playback involve a model to produce performance sequences [1] [2].

Deep Learning is the most used machine learning's algorithms [3] [4], it deals with the Artificial Neural Networks (ANN) which is inspired from the brain's structure and function. Deep Learning methods have been well developed and broadly implemented over the past few years to derive information from diverse data types [5-7]. There are many forms of architectures for Deep Learning, such as the RNN, Convolutional Neural Network (CNN), and Deep Neural Network (DNN), considering the various characteristics of input data [8] [9]. CNN and DNN are typically unable to deal with temporal input data information. Thus, RNNs dominate study fields that include sequential data, such as audio, text, and video. The two kinds of RNNs, discrete-time RNNs, and continuous-time RNNs, have different structures and functions [10] [11]. Sequence prediction RNNs are the subject of this review.

RNN is a generalized form of feedforward neural networks, improved by transferring data across time measures [12] [13]. They are a wealthy family of a new paradigm of being calculated almost randomly. In reality, the capacity to design temporal dependencies makes RNNs ideal for tasks where non-independent sequences of points consist of input and output [14]. A cyclic relation is a standard feature of the RNN architecture, allowing the RNN to modify the present state based on the state and input data that the RNN saw in the past. RNNs are capable of

transitioning between current states by using previously recorded state and current input data. The following networks, composed of normal recurrent cells (e.g., sigma cells), such as complete RNNs and selective RNNs, have had incredible success on specific issues [15] [16] [17]. Unfortunately, the above RNNs cannot link the relevant data when the data input is far apart. Suggested extended short-term memory management of strong prediction, units, or blocks of (LSTM) is part of an RNN system [18]. RNNs are created to use some forms of mechanisms of artificial memory that can assist these applications in artificial intelligence to mimic human cognition more effectively. BRNN connects two outputs of hidden layers, allowing forecasts about all possible states. Generative deep learning techniques have a more general solution than unsupervised or semi-supervised learning techniques regardless of the particular consequences of computing the models [19] [15] [16]. It was learned to forecast both positive and negative time directions concurrently. They were not widely used in unsupervised tasks since it was challenging to view the model in a probabilistic way [20] [21].

The paper is structured as follows: theoretical background in section 2 architectures of RNNs. In section 3, related work to summarize previews works is in section 4, the comparison and discussion, and finally conclusion in section 5.

## 2. BACKGROUND THEORY

RNNs are used, in particular, to address problems that are not modifiable. A neural network can understand, adjust, predict, and classify [2] [17]. The parameter capacity estimation of neural networks is robust [22]. It produces more precise outcomes than traditional predictive prediction approaches, as seen in Fig. 1.

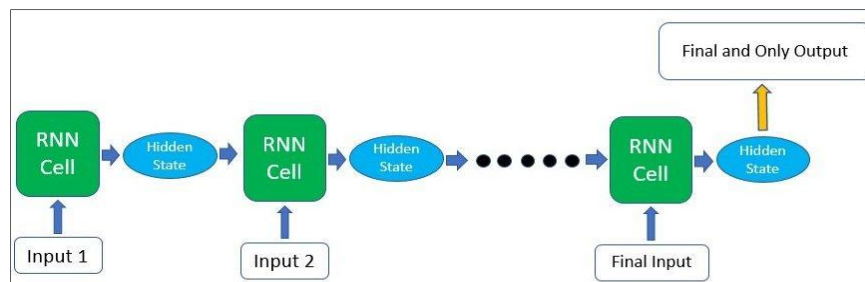


Fig. 1. Prediction of RNN [23]

At time step  $t$ , the input of the network is defined as  $x(t)$ . For example, one-hot vector  $x_1$  may be a vector that describes a word in a phrase. A hidden state:  $h(t)$  reflects the memory of the network at time  $t$ . The hidden state from the previous time step is used to determine the new hidden state. In addition to the weight mentioned above matrices, there are the weight matrices for hidden-to-hidden recurrent connections, hidden-to-output connections, hidden-to-hidden recurrent connections, and hidden-to-output connections, and all shared across time. It demonstrates the network's output. Nonlinearity is also present in  $o(t)$ , as indicated by the arrow after it.

### 2.1 RNN

RNN is a branch of Artificial Neural Network (ANN) that is widely used in the sequential knowledge applications that requires time series details. It is one of the deep learning algorithms that is utilized to solve the ordinary or temporary problems, such as text translation, voice recognition, Natural Language Processing (NLP), and image captioning. They are incorporated into popular applications such as Siri, voice search [19]. RNN, such as feedforward and CNNs, use training data to learn. They are distinguished from their 'memory' when they take knowledge from past inputs to affect the new input and performance. While traditional deep neural networks assume that the inputs and outputs of each other are independent, the success of RNN builds on the available elements of the series [24]. While

future events may indeed help to determine the success of a given sequence, these events cannot be accounted for in their predictions through nonlinear RNN [25] [26] [27]. Current neural networks are a widespread way of delivering neural feedback, enhanced by incorporating repeated borders covering the next time levels and introducing the construction with a time definition, as shown in Fig. 2.

There are three main types of RNN architectures: LSTM, BRNN, and GRU [29].

### 2.2 Long Short-Term Memory (LSTM)

(Hoch Reiter) and (Schmid Huber) suggested the LSTM cell deal with the question of "long-term dependencies." By adding a "gate" into the cell, they increased the recalling ability of the regular recurrent cell [30]. Variations consist of LSTM without a gate to forget, LSTM with a gate to forget, and LSTM with a peephole connection. A forgotten gate typically denotes LSTM with the word LSTM cell. First, the authors implement the original principle of LSTM incoming and outgoing gates [31] [32]. RNNs have trouble with short-term memory. They have difficulty transferring earlier steps to later ones if a sequence is lengthy enough [33]. For instance, RNNs may leave out essential information at the beginning of a paragraph to make predictions. LSTMs are unique sort of RNNs that can learn long-term dependencies. These types of LSTMs can retain information for lengthy durations [34] [35].

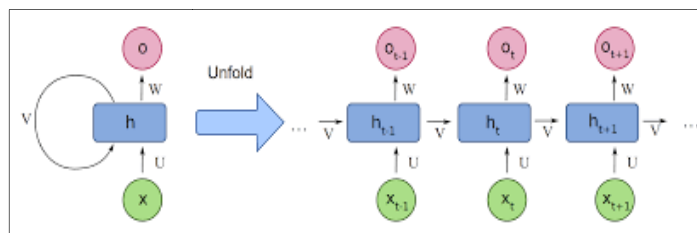


Fig. 2. Simple RNN Architecture [28]

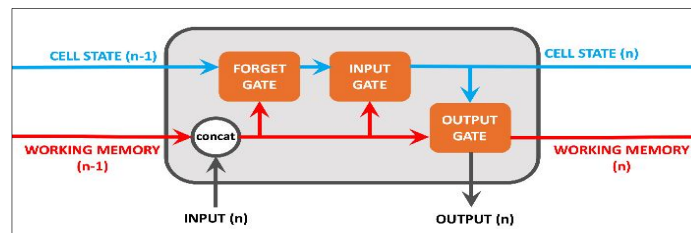


Fig. 3. The architecture of LSTM [36]

Each Unit in typical LSTM includes cell, input gate, output gate, and gate to forget. At arbitrary periods, the cell recalls values, and the three gates control the flow of knowledge into and out of the cell. Based on time-series data, LSTM networks are well adapted to classifying, analyzing, and making predictions since there may be lags of indefinite length between significant events in a time series [37]. Conventional RNNs face the vanishing gradient problem, which arises when training them. LSTMs have been created for various uses; relative insensitivity to gap duration benefits LSTM over RNNs, secret Markov models, and other sequence learning methods [38] [36]. LSTM has a three-step mechanism that indicates that the LSTM module has three gates labelled as Forget gate, input gate, an output gate, as seen in Fig. 4.

**Forget Gate:** Controls what memory knowledge to throw out, determine how many you can recall about the past.

**Update/Input Gate:** Controls whether new data is applied from the existing input to the cell state, determines how much is applied to the existing condition of this device.

**Output Gate:** determine conditional whether to produce from memory, determine which focus of the recent cell can make it to the production.

The LSTM method of preparation, such as Backward Propagation Through Time (BPTT), is a learning algorithm for teaching neural networks [40-41]. When training a neural network, they are simply optimizing the network's weights to reduce error for the previously available valid values (labels). The supervised learning algorithm is applied to identify mistakes for known names. However, it is very sluggish. For a specific time, the present data amount depends on their past data, such as sequences and sound waves. RNN has standard features and is helpful in training RNN. Besides the problems described previously, BPTT training has gradient loss and slow convergence [42]. Long-term Recurrent Network (LSTM) is a very common RNN in recent years. LSTM is one of the deep learning frameworks. The LSTM eliminates the theoretical issue of gradient depletion by utilizing gate modules. However, LSTM also requires more computer processing resources [43].

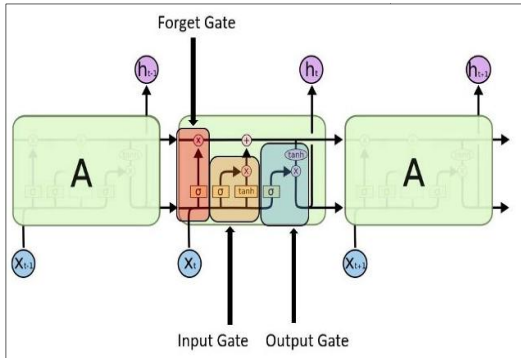


Fig. 4. Forget gate, input gate, and output gate [39]

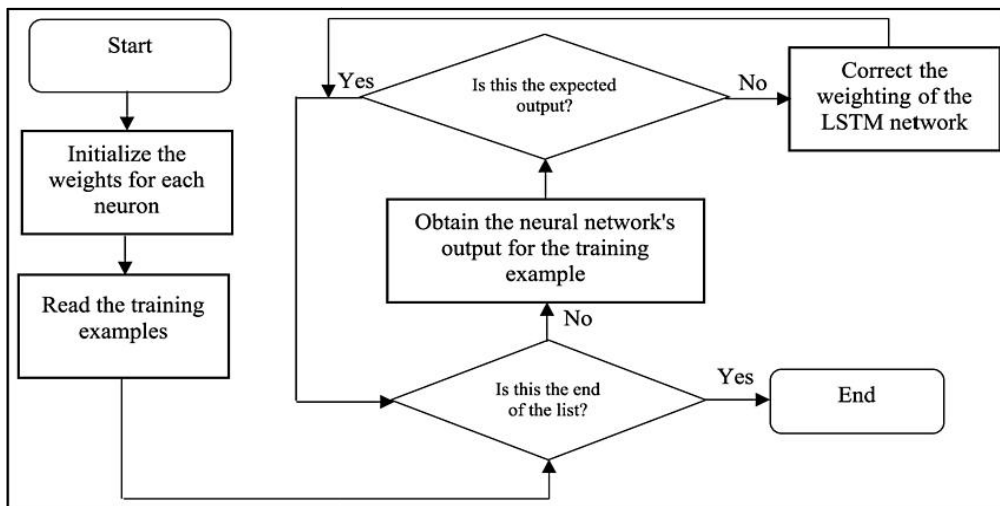


Fig. 5. Training of LSTM [44]

### 2.3 Bidirectional Recurrent Neural Network (BRNN)

There are variations of the RNN network architecture. Although unidirectional RNNs can only make conclusions about the current state from past inputs, bidirectional RNNs bring in future data to increase its precision. One of the most used RNN configurations, along with the LSTM, is the Bidirectional configuration [45]. First mentioned in [46], the recurrent neural network (BRNN) Fig. 6 RNN with two hidden layers running in opposite directions allows them to accept past and future feedback. In supervised learning methods, the generative adversarial autoencoder is more popular than semi-supervised or unsupervised learning [9]. They're trained with algorithms that are similar to RNN because the lateral neurons don't bind to each other [47]. Any additional step is required if backpropagation is necessary since improvements must be made just one time [48].

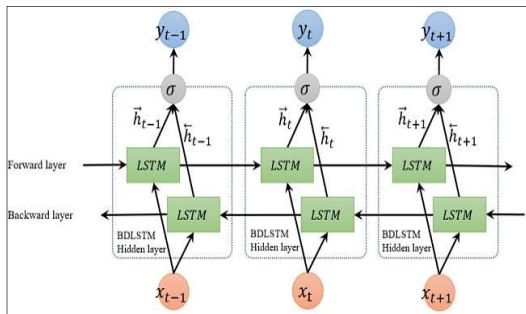


Fig. 6. The architecture of BRNN [49]

Neural Networks informal training are accessed first in the forward pass until output neurons are passed [51] [52]. The reverse takes place on the backward pass; output neurons are first stored, then forward, and backward stages are next passed [53]. The only item that is adjusted on the minute is the bar. The synaptic link designs of RNN and BRNN are identical to those observed in the human brain [47] [53]. If backpropagation is needed, an intermediate phase is required since not all input and output layers can be updated at once.

Unlike typical RNN, BRNNs are trained to predict both positive and negative pathways of time simultaneously. BRNN divides a typical RNN's neurons into two directions, one for forwarding states (positive time path) and another for backward states (negative time direction) [54]. None of these output states in

the opposite direction is connected to inputs. Input data from the possible futures of the present cycle may be used to calculate the same utility by simultaneously utilizing two-time directions [50] [55]. The reverse version of the recurrent networks that are used in bringing previous expertise forward of information. For this form of deep conceptual learning, the output layer can at the same time obtain derived from previous and future states. To improve the quantity of input information to the network, BRNNs were added [56]. There are also constraints on the RNN as the potential input knowledge cannot be obtained from the current state. BRNNs, on the other hand, do not demand that their input data be set. In addition, from the current state, their future input knowledge is available. BRNN is especially helpful when the input meaning is necessary [57]. Two separate RNNs are easily placed together by bidirectional RNNs. The input sequence is fed to one network in the usual time order, and to another network in the reverse time order. At each time point, the outputs of the two networks are typically concatenated, but there are other choices, such as summation [58] [59].

### 2.4 Gated Recurrent Unit (GRU)

A GRU seems more complicated than a conventional LSTM diagrammatically. It's also a little smoother; moreover, it trains somewhat quicker than an LSTM. Due to its relative simplicity [60]. GRUs are combined into a single update gate which acts as both an input gate and a forget gate. In reality, this suggests that forgotten cell state locations entry points will counterbalance new achievements [61] [62]. A single hidden state layer has been built that incorporates both the cell state and hidden output. The device also includes an internal, intermediate hidden state, another key distinction of the GRU. GRUs are a capable LSTM variant, and since their inception, they've been quite well-known. Although they can learn faster on tasks such as text generation or music, they often seem less efficient than their conventional LSTMs [63] due to their shortcomings in counting. The LSTM cell's learning potential is equivalent to that of the normal recurrent cell. The additional parameters, however, raise the workload of computation. The gated recurrent unit (GRU) was then implemented by [64] [65]. The details of the GRU cell layout and relations are seen in Fig. 7.

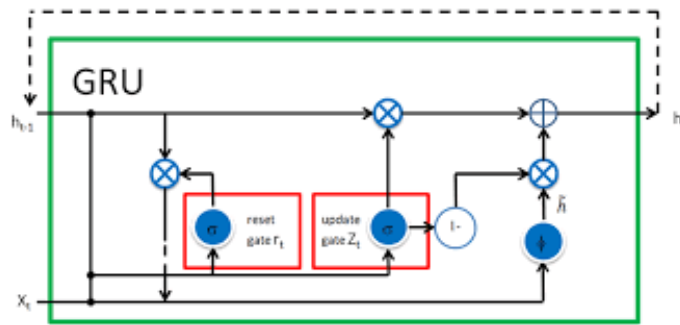


Fig. 7. The architecture of GRU [66]

The original GRU and tested the properties of three versions of GRU-1, GRU-2, and GRU-3. The findings showed that these three versions and the original GRU cell could reduce the computational cost during performance [67].

### 3. RELATED WORK

RNN is a neural sequence model that achieves state-of-the-art performance on essential data. [68] RNNs are widely used in the following domains applications: speech recognition, video tagging, machine translation, text summarization, face detection. This literature study would be dedicated to presenting the RNN, the approach of predictive development.

Li et al. [71] proposed two adaptation models for recurrent language models of the Neural Network (RNNLMs) that were suggested to capture subject effects and long-distance stimuli for automated conversational speech recognition (ASR). To adjust an RNNLM, they use a Fast Marginal Adaptation (FMA) structure. Our first model is essentially a cache model. The second model will, in theory, cause model and subject results but is tougher to train. That means that the usage of our tools is not exclusive to conversational speech recognition.

Fan et al. [62], a network traffic predictive algorithm focused on the neural network, was primarily investigated by the author. The Gated Recurrent Unit (GRU) and the RNN are used in internet traffic prediction and analysis. The computational and simulation results demonstrated the importance of this.

Jabeen et al. [71] The authors applied video retrieval based on content that uses convolution-derived genres and a recurrent network. A hundred video set of approximately 300 to 500 frames per film was utilized to

extract visual features. Genre classification using multi-class RNN depending on these visual elements, a video is named. Visual characteristics such as the face, gender, emotion, age, actions, setting, and computation are performed on objects for a query movie. The genres of a video are defined depending on the high-level components (HLFs). Initially, HLFs create a sentence followed by two simultaneous stacks of convolutional long term memory to preserve a spatial and temporal correlation of characteristics (LSTM). By comparing the genres of training images, domain videos are readily retrieved.

Armani et al. [72] proposed For IoT network protection, a Fog Computing-based network security paradigm was suggested. The suggested model adopts an RNN equipped with an improved variant of the backpropagation algorithm. The efficiency assessment findings demonstrate the efficacy of adaptive cascaded filtering utilizing the recursive neural network framework. Each network is adaptively calibrated to various parameters/hyperparameters to identify particular forms of intrusion. As a result, the model demonstrates strong sensitivity to DoS attacks that reflect one of the primary attacks that thwart IoT network creation in addition to detecting other forms of categories of attacks such as Probe, R2L, and U2R in a reasonable communication complexity as each record takes an average of 66 used to be processed. Thus, in real-time settings, the suggested framework is capable of operating properly and effectively.

Erichson et al. [73] suggested a modern recurring unit of Lipschitz that excels in various benchmark activities. Control theory claims have been used, the recurrent unit's unique configuration helps one achieve assurances of global exponential stability. The findings from

this research, in particular, inspired the symmetric-skew estimation technique for hidden-to-hidden matrices to be developed, which alleviates the issue of disappearing and bursting gradients. A model has been obtained more sensitive to input and parameter disturbances relative to other continuous-time units due to the exciting features of the system of the Lipschitz recurrent unit. The model's Hessian analysis often expresses this activity.

Nah et al. [74] suggested a strategy to improve the video deblurring recurrent network. By changing the hidden state to the goal frame iteratively, our approach more reliably eliminates blurs in the video frames. In addition, they practice our model along stochastic computing paths with a continuous function that could increase prediction accuracy. Compared to other state-of-the-art methods, our process does not consider additional parameters while also quick and effective.

Lerner et al. [75] presented a model for drug knowledge extraction developed in the French clinical texts. The objective was to discover the medicinal product's name (or type of medicinal product) in addition to a few more fields advising its administration: amount, concentration, length, state, and administration route. The role also includes the duty of extracting incidents linked to a medication order, such as a stopped drug or a drug turn. Of interest, in "Aspirin and Plavix, daily," a field may be connected to many opioid entities, such as the frequency field.

Podschwadt and Takabi [76] proposed a method that enables homomorphically encrypted information based on the CKKS system to be used by recurrent networks. In our case, feeling research on the IMDb dataset, they proposed a solution to conduct NLP tasks over encryption utilizing RNNs. Especially in comparison to the plaintext model, they can do this with no reduction of precision. So is rendered possible by adding connectivity to refresh the noise between client and server. Also, exchange network traffic for allowing word embedding to be used effectively. Our future study seeks to investigate other recurring structures such as LSTM and GRU.

Kratzer et al. [77] proposed a novel calculation system for human motion in the attendance of environmental targets. They demonstrated that within a trajectory optimization system, a

recurrent government system could be modified for use. It strengthens forecasts and allows for environmental structures that are not apparent. In addition, they demonstrated an initial trial about how the methodology can be used for shared human-robot preparation.

Datta et al. [78] had computer translation programs built to understand the difficulties created by this. To promote machine translation, engineers of many respectable organizations such as Google LLC have sought to implement applications to enable machine translations utilizing machine learning algorithms such as the Artificial Neural Network (ANN). In this respect, many neural machine translations have been done, but, on the other hand, the RNN (RNN) has not developed much in this area. In this job, to understand the advantages of RNN over ANN, authors have attempted to introduce RNN into the area of computer translations. The findings demonstrate how RNN is capable of doing computer translations with sufficient precision.

Manchev and Spratling [79] proposed implemented a new training algorithm for recurrent networks, target propagation over time (TPPT), that outperforms traditional time-backpropagation (BPTT) on four of the research issues. On four simulated time lag tasks, the proposed algorithm is initially checked and compared to BPTT, and its output is also calculated using the sequential MNIST data collection. Moreover, it allows for different nonlinearities as TPPT uses goal propagation which could theoretically alleviate credit assignment in more complicated recurrent architectures.

Karevan and Suykens [80] suggested to achieve a data-driven forecasting model for the usage of weather forecasting, LSTM was used. In addition, Transudative LSTM (T-LSTM) was suggested to manipulate local data in time-series prediction. For the dilemma of regression, a quadratic cost process is chosen. The objective function is localized by considering a weighted equation cost function, at which stage the samples have greater weights in the neighbourhood of the test point. Two weighting systems are investigated based on the cosine correlation between the training samples and the evaluation stage. The tests are performed over two different amounts of time each year to determine the efficiency of the proposed process under different weather

conditions. The findings indicate that in the prediction mission, T-LSTM results in better efficiency.

Fabien et al. [81] suggested obtaining insights into the nature of NN for general velocity inversion. They suggested researching the simplistic problem of 1D velocity estimation. They have shown that one can depict seismic data appropriate for automatic velocity analysis using a deep convolutional NN. The arms and interval velocities were calculated based on this representation with recurrent NNs, for dynamic 1D momentum models with high-velocity parallels and thin layers. That indicates that a data-driven method utilizing NNs can handle realistically dynamic environments to simplify velocity analysis and implies the plausibility of automating current velocity method workflows used for subsurface imaging.

Gao et al. [82] proposed two approaches focused on DNN that have been suggested to overcome goal motion instability. Specifically, it suggested using RNN-based networks with deep frameworks to estimate the target systems from noisy measurements, considering that the RNN operates sequentially and may infer a conditional density with its repeating modules. While the real-time application of the proposed solutions was shown to be practicable, it was also considered a way of reducing the computing load at the cost of performance. Numerical findings have shown that the proposed methods have the ability, in terms of statistical estimation precision, to transcend the limitations of conventional model-based methods, especially during the early stages of inadequate estimation data.

Shewalkar et al. [83] suggested analyzed and contrasted their results on a decreased TED-LIUM speech data collection with RNN, LSTM, and GRU. Two different architectures were assessed; a 500-node network and a 1,000-node network in each layer. WER, failure, mean edit range and the runtime were the assessment tools used. The findings reveal that the LSTM and GRU WER values are similar (LSTM performs marginally higher than GRU), but the LSTM runtime is more significant than GRU. Therefore, as it returned excellent WER values within a reasonable runtime, the decrease TED-LIUM speech data collection suggests using GRU.

Ma et al. [84] proposed a novel LSTM HSI classification system, where unlabeled data is

well-exploited to create serial features from a single HSI. Corresponding pixels obtained from the whole image are used to create the respective serial features instead of spectral features as the serial data structure of LSTM. Specifically, the resemblance between a goal pixel and other such pixels in the picture is treated when creating a sequential function. Two similarity-measuring approaches-pixel-matching and frame introduced here to help represent the similarity of two pixels, where individual spectral characteristics are used in pixel-matching-based systems, and all spatial and spectral data are used in block-matching-based systems.

Gao et al. [85] indicated that RNN-related predictors have the prediction value over the next century of TBM operational parameters based on in-situ monitoring results. The predictors consist of four sections: the input layer, the RNN layer, the intertwined layers, and the output layer. As one of three modes of RNNs, the RNN layer can be precisely set: regular RNNs, LSTM, and GRU. Linked numerical tests validate the validity of the predictor variables for six main functioning TBM parameters, including torque (T), velocity (V), thrust (F), upper left chamber pressure (PCTL), lower left chamber pressure (PCBL), and lower proper chamber pressure (PCBR) based on actual proof (PCBR). Experimental results indicate that 1) the RNN-based predictors function well for these operational conditions, 2) the RNN-based gate operating predictors (e.g., LSTM and GRU) surpass the traditional RNN-based predictors, and 3) the GRU-based predictor, which performs better than the LSTM-based one, with simpler gate operations, for five (out of six) operating conditions. In comparison, the classical one was an embrace.

Lemayian et al. [86] suggested existing mMIMO channel prediction schemes focused on RNN-based and suggested a low-difficulty, low-cost channel forecasting scheme. The findings showed that RNN-based CSI prediction is the perfect technology to increase the efficiency of mMIMO systems by reducing sophistication and increasing precision. IMO is a technology that has demonstrated tremendous potential to enhance wireless connectivity in the future. Therefore, they plan to continue this work by improving the estimation of mMIMO CSI using various RNN network configurations using LSTMS or GRUs and enhancing the planned predictor's effectiveness.



Uddin et al. [87] suggested a deep RNN-dependent behaviour recognition device based on fusion data from the wear-ability body sensor. After removing the data from the sensors, the multimodal data is merged and accompanied by KPCA to render it stable. Then, a deep RNN is trained for all the operations, later used to evaluate the results. They performed extensive studies on three publicly accessible databases to verify and compare our proposed method by achieving 0.99 accuracy, recall, F1-score, and broad datasets.

Yin et al. [88] suggested creating compact RNN models utilizing decomposition of Hierarchical Tucker (HT). HT decomposition provides the decomposed RNN models with a clear hierarchical framework, which is very helpful and essential for improving the capacity for representation. Meanwhile, HT decomposition promises more significant cost savings in storage and computation than the conventional tensor decomposition methods for RNN compression. The results of the experiments and neuron-computational models explicitly demonstrated that our proposed HT-based LSTM outperforms the state-of-the-art methods in both speed and test accuracy over multiple datasets.

Li et al. [89] proposed a new prediction model of protein-RNA binding based on deep neural networks has been created. The approach uses LSTM to remove dependence features and Dense Net to reuse RNA functions in each layer to acquire more powerful knowledge that will enhance precision. This technique will go towards offering a biological justification for the binding of protein-RNA. The paper also considers that in forecasting RNA binding, the pairwise likelihood features of RNA presented in this paper are indeed beneficial. Any of the protein-RNA binding experimental results did not work well in our model. That is because there were just a few overlapping proteins between the in vitro and in vitro studies. Since noise or the internal climate influences in vitro binding, RNA transcripts often require thousands of nucleotides, and the mechanism of transcription is dynamic and unpredictable. It is often hard to extract correctly in vitro predictions with the RNA pairwise likelihood function added.

Tang et al. [90] suggested the short-term power load forecast is the base for grid forecasting and

decision making. To forecast power load, a multi-layer bidirectional RNN centred on LSTM and GRU is proposed. The experimental outcome indicates that the approach proposed is superior to the competition winner in terms of the quality of the European Intelligent Higher Recognition competition data prediction. The authors estimate and evaluate the seasonal load separately with LSTM, help vector regression, and backpropagation models.

Chen et al. [21] proposed the BRIM, which separates the status neurons of the standard RNN in two sections proposed long-term bidirectional prediction (LSM) model: forward states (using experience of the background of the electricity price) directed at the processing of data in a positive direction or forward states (using future price details accessible on connected markets). In addition, it is fair to integrate and take advantage of the effect of adjacent markets on the precision of energy price forecasts. The integrated market's potential energy rates are used both as input functions for forwarding and backward LSTM. In improving predictive precision, the experimental findings indicate the supremacy of the proposed BRIM system. According to the report, BRIM statistically substantially outperforms other systems.

Bothe et al. [91] proposed a bidirectional RNN (Utt-Att-BiRNN) utterance-level attention-based model to examine the significance of defining the present one, past utterances. In our configuration, the input collection of current and preceding utterances is given to the BiRNN. Our model succeeds previous models that use as a backdrop on the utilized corpus just preceding utterances. Another contribution of our research is discovering the amount of data in each phrase to classify the corresponding one and demonstrating that context-based learning increases efficiency and achieves greater faith in the understanding of dialogue actions. They identified that the nearest preceding utterances contribute to a greater degree when classifying short utterances. The model produces a state-of-the-art outcome of around 77% accuracy on the SwDA corpus, utilizing only previous utterances in the background.

Cui et al. [92] proposed a deep-stacked neural network architecture for bidirectional and unidirectional LSTM (SBULSTM), which considers both forward and backward data dependencies from time series to Predict the

**Table 1. Summary of the prediction based on deep RNNs**

Reference	Year	Algorithm	Data Sets	Objectives	Results	Accuracy
Li et al.	2018	FMA for RNNLM, Conversational Cache Model, Adaptation and Rescoring Pipeline, and DNN	Switchboard (SWBD), Spanish, and Egyptian	Adapting an RNN Language Models (RNNLMs).	Two adaptation models to capture subject effects and long-distance stimuli.	57.5% 71.2% 47.0%
Fan et al.	2019	Gated Recurrent Unit (GRU)	Computation and simulations.	Network traffic prediction model focused on neural networks	Examining the network traffic Prediction model. The findings of numerical and simulation showed the efficacy of	MSE 1.011% NMSE 0.972% MARE 1.171%
Jabeen et al.	2019	High-level features (HLFs).	CoNLL-2003 and OntoNotes 5.0.	Used Parallel Multi-Class RNN Based on Video Description	Images that might have scenes or items may even be monitored without humans.	88.13%
Nah et al.	2019	Intra-Frame Iterations(denote as IFI-RNN)	GOPRO	Adapting the secret states passed between past frames to a frame being analyzed.	They showed that our technique demonstrates government video deblurring efficiency when running at the real-time level.	29.16% 29.97%
Almiani et al.	2020	Mathew correlation and Cohen's Kappa	NSL-KDD	Model of intrusion detection focused on Fog computing for IoT network protection	Exhibiting significant amounts of instances where there is DoS assault.	92.42%
Erichson et al.	2020	Hessian matrices	TIMIT and Penn Tree Bank (PTB)	Defining the evolution of the secret state with two parts: a linear portion well-understood plus a nonlinearity of Lipschitz	The Lipschitz recurrent unit is more stable than other continuous-time RNNs in input and parameter disruptions.	97.2% 97.5%
Lerner et al.	2020	Seq-BiLSTMs And BiLSTMs	APmed dataset	Testing the RNNG in clinical tests for drug knowledge extraction.	Detecting entities is lower than the baseline BiLSTM, with RNNG weaker than the backdrop Bil STM.	88.5%

Reference	Year	Algorithm	Data Sets	Objectives	Results	Accuracy
Podschwadt and Takabi.	2020	SMC,DP,and HE	IMDb	To build private information RNNs for natural language processing activities, homomorphic encryption was used.	Our findings suggest that RNNs can run without compromising precision and fair efficiency, and contact expense over encrypted data.	86.4%
Kratzer et al.	2020	Using a gradient-based optimization algorithm to optimize the motion	Real motion Data.	In the presence of environmental targets, a novel prediction method for the human motion was used.	The consequence is that it is possible to modify a state-of-the-art RNN architecture within a trajectory optimization system...	10%
Datta.	2020	Google NMT	IBM ViaVoice software.	In the field of machine translations, take RNN.	The bidirectional Recurrent Neural Network with algebraic equations provides high accuracy.	41.99 %
Manchev and Spratling	2020	BPTT and TPTT	MNIST	Introduced a new algorithm to train recurrent networks	An always there for the problems of artificial classics has been discovered, and the MNIST pixel sequence problem is equally correct.	54.20% 54.75%

network-wide speed of traffic. An LSTM Bidirectional sheet (BDLSTM) captures spatial characteristics and temporal bidirectional dependencies from historical records. This is the first time BDLSTMs have got the best of our expertise and applied it as building blocks for a deep architecture model to calculate the backward dependence of traffic results for prediction. The suggested model can accommodate missed input data values by using a mechanism for masking. In addition, this scalable model would predict traffic speed on all highways and dynamic urban traffic with networks. Equivalences with other state-of-the-art symbols in terms of precision and reliability, the proposed SBU-LSTM neural network has made excellent predictions for the actual traffic in the network.

#### 4. COMPARISON AND DISCUSSION

This study's latest evaluation illustrates how RNN algorithms may be applied in several branches of science. Although the effective use of machine learning as a powerful method in deep RNN has been accepted, this method has been given the green light in terms of prediction-based deep RNN. Many people think this algorithm has a crucial benefit in most of these studies. A general overview of these research findings compiled in the past year (2018, 2019, and 2020) is given in comparison Table1.

They observed consistent WER reductions. This indicates that the approaches are not restricted to conversational. The prediction and interpretation of network traffic signals, the Gated Recurrent Unit (GRU), and the RNN model are integrated with the potential to match nonlinear and multi-dimensional functions precisely. However, centred on the high-level features (HLFs), HLFs are used to produce a phrase preceded by two simultaneous stacks of Convolutionary Long-Term Memory (LSTM) videos from the database, obtained by comparing the training video genres. This retrieval can further be extended by tracking some specific labels. Reveals how the writers posed an automatic intrusion detection framework for Fog protection against cyber-attacks. The five layers of hierarchical RNNs that are used for Fog computing protection are pretty appealing. Moreover, it evaluates and compare the Lipschitz RNN performance with other state-of-the-art models. The main task of the model is to order and permuted pixel-by-

pixel MNIST classification and audio data using the TIMIT dataset. All of these tasks require that the recurrent unit learns long-term dependencies. Also, a researcher developed a single model that can be trained to handle various internal recurrence paths. Therefore, they show the superiority of the proposed model over the current state-of-the-art methods in both deblurring accuracy and computational efficiency. RNN's partnership success can be clarified by the model design, which offers a workaround between distant sections of the word, and the mutual modelling target that helps RNN understand richer representations.

On the other hand, the researchers propose an approach that combines RNNs, Elman networks, and homomorphic encryption to infer encrypted data. Results demonstrate that it can run RNNs over encrypted Data without sacrificing accuracy, good performance, and communication cost.

#### 5. CONCLUSION

This paper aimed to review many papers to cover recent innovations in the prediction based on RNN. Recurrent networks are particularly fascinating because they seem capable of overcoming many of the severe constraints that conventional machine learning methods normally impose on data. The presumption of independence between successive examples is broken with recurrent networks, assuming that fixed-dimension feedback is broken. Yet, on several tasks, RNN models perform competitive level with or outperform state of the art.

#### COMPETING INTERESTS

Authors have declared that no competing interests exist.

#### REFERENCES

1. Abdulqader D, Abdulazeez A, Zeebaree D. Machine Learning Supervised Algorithms of Gene Selection: A Review; 2020.
2. Rahman A, Sri Kumar V, Smith AD. "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," Appl. Energy. 2018;212:372–385.
3. Maulud D, Abdulazeez AM. A Review on Linear Regression Comprehensive in Machine Learning, J. Appl. Sci. Technol. Trends. 2020;1(4):140–147.

4. Kareem O, Mohsin Abdulazeez A, Zeebaree D. Skin Lesions Classification Using Deep Learning Techniques: Review. 2021;1–22.  
DOI: 10.9734/AJRCOS/2021/v9i130210.
5. LeCun Y, Bengio Y, Hinton G. Deep learning," nature. 2015;521(7553):436–444.
6. Yan LC, Yoshua B, Geoffrey H. Deep learning," nature. 2015;521(7553):436–444.
7. Barwary M, Mohsin Abdulazeez A. Impact of Deep Learning on Transfer Learning : A Review; 2021.  
DOI: 10.5281/zenodo.4559668.
8. Mohammed S, Mohsin Abdulazeez A. Deep Convolution Neural Network for Facial Expression Recognition Deep Convolution Neural Network for Facial Expression Recognition Deep Convolution Neural Network for Facial Expression Recognition ---Palarch's Journal Of Archaeology Of Egypt/Egyptology. ISSN 1567 -214x Keywords-facial expression recognition, principle component analysis, convolutional neural network; 2021.
9. Mohsin Abdulazeez A, Zebari D, Jijo B. Machine Learning Classifiers Based Classification For IRIS Recognition," Qubahan Acad. J. 2021;1.  
DOI: 10.48161/qaj.v1n2a48.
10. Zeebaree DQ, Haron H, Abdulazeez AM. Gene Selection and Classification of Microarray Data Using Convolutional Neural Network. 2018;145–150.  
DOI: 10.1109/ICOASE.2018.8548836.
11. Hassan R, Mohsin Abdulazeez A. Deep Learning Convolutional Neural Network for Face Recognition: A Review; 2021.  
DOI: 10.5281/zenodo.4471013.
12. Abas Hasan D, Mohsin Abdulazeez A. "A Modified Convolutional Neural Networks Model for Medical Image Segmentation," Test Eng. Manag. 2020;83:16798–16808.
13. Saeed J, Mohsin Abdulazeez A. "Facial Beauty Prediction and Analysis Based on Deep Convolutional Neural Network: A Review, J. Soft Comput. Data Min. 2021;2.  
DOI: 10.30880/jscdm.2021.02.01.001.
14. Saad Hikmat H, Adnan Mohsin A. Comparison of Optimization Techniques Based on Gradient Descent Algorithm: A Review, PalArchs J. Archaeol. Egypt Egyptol. 2021;18(4):2715–2743.
15. Sadeeq H, Abdulazeez A, Kako N, Abraham A. "A Novel Hybrid Bird Mating Optimizer with Differential Evolution for Engineering Design Optimization Problems. 2018;522–534.
16. Mohsin Abdulazeez A. Journal of Soft Computing and Data Mining Evaluating Data Mining Classification Methods Performance in Internet of Things Applications; 2020.  
DOI: 10.30880/jscdm.2020.01.02.002.
17. Mohammed I, Adeen N, Abdulazeez A, Zeebaree D. "Systematic Review of Unsupervised Genomic Clustering Algorithms Techniques for High Dimensional Datasets; 2020.
18. Zebari R, Abdulazeez A, Zeebaree D, Zebari D, Saeed J. A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," J. Appl. Sci. Technol. Trends. 2020;1(2):56–70.  
DOI: 10.38094/jastt1224.
19. Zeebaree DQ, Haron H, Abdulazeez AM, Zebari DA. "Machine learning and Region Growing for Breast Cancer Segmentation," in International Conference on Advanced Science and Engineering (ICOASE). 2019;88–93.
20. Zebari D, Abdulazeez A, Zeebaree D, Salih M. A Fusion Scheme of Texture Features for COVID-19 Detection of CT Scan Images; 2020.
21. Ordóñez FJ, Roggen D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," Sensors. 2016;16(1):115.
22. Najat N, Mohsin Abdulazeez A. Gene clustering with partition around mediods algorithm based on weighted and normalized mahalanobis distance. 2017;145.  
DOI: 10.1109/ICIIBMS.2017.8279707.
23. Eesa A, Mohsin Abdulazeez A, Orman Z. A New DIDS Design Based on a Combination Feature Selection Approach; 2015.
24. Abas Hasan D, Mohsin Abdulazeez A. A Modified Convolutional Neural Networks Model for Medical Image Segmentation," Test Eng. Manag. 2020;83:16798–16808.
25. Ma F, Chitta R, Zhou J, You Q, Sun T, Gao J. Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks," in Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 2017;1903–1911.
26. Yan Z, Lécuyer E, Blanchette M, "Prediction of mRNA subcellular

- localization using deep recurrent neural networks," *Bioinformatics*. 2019;35(14):i333–i342.
27. Bartlett Z, Han L, Nguyen TT, Johnson P. Prediction of road traffic flow based on deep recurrent neural networks," in 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). 2019;102–109.
  28. Selvin S, Vinayakumar R, Gopalakrishnan EA, Menon VK, Soman KP. "Stock price prediction using LSTM, RNN and CNN-sliding window model," in 2017 international conference on advances in computing, communications and informatics (icacci). 2017;1643–1647.
  29. Mehyadin A, Mohsin Abdulazeez A, Abas Hasan D, Saeed J. Birds Sound Classification Based on Machine Learning Algorithms," *Asian J. Res. Comput. Sci.* 2021;1–11.  
DOI: 10.9734/AJRCOS/2021/v9i430227.
  30. Charbuty B, Abdulazeez A. "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*. 2021;2(01):20–28.  
DOI: 10.38094/jastt20165.
  31. Graves A, Mohamed A, Hinton G. "Speech recognition with deep recurrent neural networks," in 2013 IEEE international conference on acoustics, speech and signal processing, 2013;6645–6649.
  32. Chan W, Lane I. Deep recurrent neural networks for acoustic modelling," *ArXiv Prepr. ArXiv150401482*; 2015.
  33. Kazheen I, Mohsin Abdulazeez A. Deep Learning Convolutional Neural Network for Speech Recognition: A Review; 2021.  
DOI: 10.5281/zenodo.4475361.
  34. Mikolov T, Joulin A, Chopra S, Mathieu M, Ranzato M. Learning longer memory in recurrent neural networks," *ArXiv Prepr. ArXiv14127753*; 2014.
  35. Ren H, Wang W, Liu C. Recognizing online handwritten Chinese characters using RNNs with new computing architectures, *Pattern Recognit.* 2019;93:179–192.
  36. Hochreiter S, Schmidhuber J. LSTM can solve hard long time lag problems in Advances in neural information processing systems. 1997;473–479.
  37. Hochreiter S, Schmidhuber J. "Long short-term memory," *Neural Comput.* 1997;9(8):1735–1780.
  38. Graves A. Long short-term memory," in *Supervised sequence labelling with recurrent neural networks*, Springer. 2012;37–45.
  39. Sak H, Senior AW, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling; 2014.
  40. Rashid H, Mohsin Abdulazeez A, Zebari D. Data Mining Classification Techniques for Diabetes Prediction," *Qubahan Acad. J.* 2021;1:17.  
DOI: 10.48161/qaj.v1n2a55.
  41. Zhu X, Sobihani P, Guo H. Long short-term memory over recursive structures," in *International Conference on Machine Learning*. 2015;1604–1612.
  42. Bansal T, Belanger D, McCallum A. Ask the gru: Multi-task learning for deep text recommendations," in *Proceedings of the 10th ACM Conference on Recommender Systems*. 2016;107–114.
  43. Althé F, de La Fortelle A. "An LSTM network for highway trajectory prediction," in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). 2017;353–359.
  44. Li Y, Zhu Z, Kong D, Han H, Zhao Y. "EA-LSTM: Evolutionary attention-based LSTM for time series prediction," *Knowl.-Based Syst.* 2019;181:104785.
  45. Wang Y, Huang M, Zhu X, Zhao L. "Attention-based LSTM for aspect-level sentiment classification," in *Proceedings of the 2016 conference on empirical methods in natural language processing*. 2016;606–615.
  46. You Y, Hseu J, Ying C, Demmel J, Keutzer K, Hsieh CJ. Large-batch training for LSTM and beyond," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2019;1–16.
  47. Ranzato M, Chopra S, Auli M, Zaremba W. Sequence level training with recurrent neural networks, *ArXiv Prepr. ArXiv151106732*; 2015.
  48. Tan T, et al. Speaker-aware training of LSTM-RNNs for acoustic modelling," in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016;5280–5284.
  49. Mehyadin A, Mohsin Abdulazeez A. "Classification Based on Semi-Supervised

- Learning: A Review," Iraqi J. Comput. Inform. 2021;47.
50. Ergen T, Kozat SS. Online training of LSTM networks in distributed systems for variable length data sequences," IEEE Trans. Neural Netw. Learn. Syst. 2017;29(10):5159–5165.
  51. Abdulkareem NM, Mohsin Abdulazeez A, Qader Zeebaree D, Hasan DA. "COVID-19 World Vaccination Progress Using Machine Learning Classification Algorithms," Qubahan Acad. J. 2021;1(2): 100–105.  
DOI: 10.48161/qaj.v1n2a53.
  52. Karanov B, Lavery D, Bayvel P, Schmalen L. End-to-end optimized transmission over dispersive intensity-modulated channels using bidirectional recurrent neural networks," Opt. Express. 2019;27(14):19650–19663.
  53. Fan Y, Qian Y, Xie FL, Soong FK. TTS synthesis with bidirectional LSTM based recurrent neural networks; 2014.
  54. Qiao K, et al. "Category decoding of visual stimuli from human brain activity using a bidirectional recurrent neural network to simulate bidirectional information flows in human visual cortices," Front. Neurosci. 2019;13.
  55. Kim J, El-Khamy M, Lee J. Residual LSTM: Design of a deep recurrent architecture for distant speech recognition," ArXiv Prepr. ArXiv170103360; 2017.
  56. Abdulrahman L, Mohsin Abdulazeez A, Abas Hasan D. COVID-19 World Vaccine Adverse Reactions Based on Machine Learning Clustering Algorithm; 2021.
  57. Chang LC, Chen PA, Chang FJ. "Reinforced two-step-ahead weight adjustment technique for online training of recurrent neural networks," IEEE Trans. Neural Netw. Learn. Syst. 2012;23(8): 1269–1278.
  58. Ibrahim I, Abdulazeez A. The Role of Machine Learning Algorithms for Diagnosing Diseases," J. Appl. Sci. Technol. Trends. 2021;2(01):10–19.  
DOI: 10.38094/jastt20179.
  59. Sundermeyer M, Alkhoul T, Wuebker J, Ney H. "Translation modeling with bidirectional recurrent neural networks," in Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014;14–25.
  60. Na T, Ko JH, Kung J, Mukhopadhyay S. On-chip training of recurrent neural networks with limited numerical precision," in 2017 International Joint Conference on Neural Networks (IJCNN). 2017;3716–3723.
  61. Zeebaree DQ, Haron H, Abdulazeez AM, Zebari DA. Trainable Model Based on New Uniform LBP Feature to Identify the Risk of the Breast Cancer. 2019;106–111.  
DOI: 10.1109/ICOASE.2019.8723827.
  62. Rashid H, Mohsin Abdulazeez A, Abas Hasan D. Detection of Diabetic Retinopathy Based on Convolutional Neural Networks: A Review," Asian J. Res. Comput. Sci. 2021;1–15.  
DOI: 10.9734/AJRCOS/2021/v8i330200.
  63. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Neural Netw. 2005;18(5–6):602–610.
  64. Sun L, Kang S, Li K, Meng H. Voice conversion using deep bidirectional long short-term memory based recurrent neural networks, in 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP). 2015;4869–4873.
  65. Salah N, Mohsin Abdulazeez A, Zeebaree D, Abas Hasan D. Medical Images Breast Cancer Segmentation Based on K-Means Clustering Algorithm: A Review," Asian J. Res. Comput. Sci., pp. 23–38, 24 2021,  
DOI: 10.9734/AJRCOS/2021/v9i130212.
  66. Pollastri G, Przybylski D, Rost B, Baldi P. Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles," Proteins Struct. Funct. Bioinforma. 2002;47(2):228–235.
  67. Farsad N, Goldsmith A. Sliding bidirectional recurrent neural networks for sequence detection in communication systems," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018;2331–2335.
  68. Dey R, Salemt FM. Gate-variants of gated recurrent unit (GRU) neural networks," in 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). 2017;1597–1600.
  69. Kareem F, Mohsin Abdulazeez A, Abas Hasan D. Predicting Weather Forecasting State Based on Data Mining Classification Algorithms," Asian J. Res. Comput. Sci. 2021;9:13–24.  
DOI: 10.9734/AJRCOS/2021/v9i330222.

70. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling," ArXiv Prepr. ArXiv14123555; 2014.
71. Kareem O, Al-Sulaifanie A, Abas Hasan D, Ahmed D. Segmenting and Classifying the Brain Tumor from MRI Medical Images Based on Machine Learning Algorithms: A Review," Asian J. Res. Comput. Sci. 2021;10:51–60.  
DOI: 10.9734/AJRCOS/2021/v10i230239.
72. Rana R. Gated recurrent unit (GRU) for emotion classification from noisy speech," ArXiv Prepr. ArXiv161207778; 2016.
73. Tang Y, Huang Y, Wu Z, Meng H, Xu M, Cai L. Question detection from acoustic features using recurrent neural network with gated recurrent unit," in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2016;6125–6129.
74. Wang Y, Liao W, Chang Y. "Gated recurrent unit network-based short-term photovoltaic forecasting," Energies. 2018;11(8):2163.
75. Chen J, Jing H, Chang Y, Liu Q. Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process, Reliab. Eng. Syst. Saf. 2019;185:372–382.
76. Dyer C, Kuncoro A, Ballesteros M, Smith NA. Recurrent neural network grammars," ArXiv Prepr. ArXiv160207776; 2016.
77. Li K, Xu H, Wang Y, Povey D, Khudanpur S. Recurrent Neural Network Language Model Adaptation for Conversational Speech Recognition., in Interspeech. 2018; 3373–3377.
78. Fan J, Mu D. Research on Network Traffic Prediction Model Based on Neural Network," in 2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China, Sep. 2019;554–557.  
DOI:10.1109/ICISCAE48440.2019.221694.
79. Jabeen S, Khan G, Naveed H, Khan Z, Khan UG. Video Retrieval System Using Parallel Multi-Class Recurrent Neural Network Based on Video Description," in 2018 14th International Conference on Emerging Technologies (ICET), Islamabad. 2018;1–6.  
DOI: 10.1109/ICET.2018.8603598.
80. Almiani M. Deep recurrent neural network for IoT intrusion detection system," Simul. Model. Pract. Theory. 2020;20.
81. Erichson NB, Azencot O, Queiruga A, Hodgkinson L, Mahoney MW. Lipschitz Recurrent Neural Networks," ArXiv200612070 Cs Math Stat; 2020, Accessed: Jan. 06, 2021. [Online]. Available: <http://arxiv.org/abs/2006.12070>
82. Nah S, Son S, Lee KM. "Recurrent Neural Networks With Intra-Frame Iterations for Video Deblurring," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA. 2019;8094–8103.  
DOI: 10.1109/CVPR.2019.00829.
83. Lerner I, Jouffroy J, Burgun A, Neuraz A. Learning the grammar of prescription: recurrent neural network grammars for medication information extraction in clinical texts," ArXiv Prepr. ArXiv200411622; 2020.
84. Kratzer P, Toussaint M, Mainprice J. Prediction of Human Full-Body Movements with Motion Optimization and Recurrent Neural Networks," ArXiv191001843 Cs; 2020, Accessed: Jan. 06, 2021. [Online]. Available: <http://arxiv.org/abs/1910.01843>
85. Datta D, David PE, Mittal D, Jain A. Neural Machine Translation using Recurrent Neural Network," Int. J. Eng. Adv. Technol. 2020;9(4):1395–1400.
86. Uddin MZ, Hassan MM, Alsanad A, Savaglio C. A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare," Inf. Fusion. 2020;55:105–115.
87. Karevan Z, Suykens JAK. Transductive LSTM for time-series prediction: An application to weather forecasting," Neural Netw. 2020;125:1–9.  
DOI: 10.1016/j.neunet.2019.12.030.
88. Fabien-Ouellet G, Sarkar R. Seismic velocity estimation: A deep recurrent neural-network approach," GEOPHYSICS. 2020;85(1):U21–U29.  
DOI: 10.1190/geo2018-0786.1.
89. Gao C, Yan J, Zhou S, Varshney PK, Liu H. Long short-term memory-based deep recurrent neural networks for target tracking," Inf. Sci. 2019;502:279–296.  
DOI: 10.1016/j.ins.2019.06.039.
90. Shewalkar A, Nyavanandi D, Ludwig SA. Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU, J. Artif. Intell. Soft Comput. Res. 2019;9(4):235–245.



91. Ma A, Filippi AM, Wang Z, Yin Z. Hyperspectral image classification using similarity measurements-based deep recurrent neural networks," Remote Sens. 2019;11(2):194.
92. Gao X, Shi M, Song X, Zhang C, Zhang H. Recurrent neural networks for real-time prediction of TBM operating parameters," Autom. Constr. 2019;98:225–235.

© 2021 Zeebaree et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*

*The peer review history for this paper can be accessed here:*  
<https://www.sdiarticle4.com/review-history/72108>