



DeTrAs: deep learning-based healthcare framework for IoT-based assistance of Alzheimer patients

Sumit Sharma¹ · Rajan Kumar Dudeja¹ · Gagangeet Singh Aujla² · Rasmeet Singh Bali¹ · Neeraj Kumar³

Received: 1 July 2020 / Accepted: 2 September 2020
© The Author(s) 2020

Abstract

Healthcare 4.0 paradigm aims at realization of data-driven and patient-centric health systems wherein advanced sensors can be deployed to provide personalized assistance. Hence, extreme mentally affected patients from diseases like Alzheimer can be assisted using sophisticated algorithms and enabling technologies. Motivated from this fact, in this paper, DeTrAs: Deep Learning-based Internet of Health Framework for the Assistance of Alzheimer Patients is proposed. DeTrAs works in three phases: (1) A recurrent neural network-based Alzheimer prediction scheme is proposed which uses sensory movement data, (2) an ensemble approach for abnormality tracking for Alzheimer patients is designed which comprises two parts: (a) convolutional neural network-based emotion detection scheme and (b) timestamp window-based natural language processing scheme, and (3) an IoT-based assistance mechanism for the Alzheimer patients is also presented. The evaluation of DeTrAs depicts almost 10–20% improvement in terms of accuracy in contrast to the different existing machine learning algorithms.

Keywords Alzheimer disease · Convolution neural network · Internet of Health · Naive Bayes · Recurrent neural network

1 Introduction

The transition of conventional healthcare systems to a data-driven and patient-centric healthcare 4.0 has initiated a pragmatic change in the health statistics [1]. The adaption of cutting-edge technologies (Internet of things, body area networks) which are driven by sophisticated data-driven algorithms (machine learning, deep learning) has supported this change in healthcare systems equipped with smart devices (wearable devices, sensors, medical gadgets) [2]. This transition has lowered the fatality rate of populations and has increased average life expectancy. The healthcare 4.0 ecosystem works largely on two foundations: (1) patients (physical world) and (2) cloud or edge/fog-enabled

algorithms and autonomous systems (virtual world) [3]. This modern healthcare system relies heavily on the cross-organizational services which tend to promote personalization and individual healthcare assistance and support using big data analytic [4]. The enabling technologies, i.e., data analytics and recommender systems, have a vast research potential in context of healthcare systems. Due to this transformation, the personalized recommendations can be provided to the patients suffering from various diseases using the enabling technologies. These recommendations consider different verticals of diseases such as staging, severity, risks involved and assistance possible. Thereafter, a balanced decision can be drawn about the possible treatments, precautions and assistance for the specific diseases [5].

Healthcare sector could be further bifurcated into physical and mental healthcare. Past decade has witnessed a shift of the healthcare industry from physical to mental healthcare [6]. Alzheimer is one of the most prevalent diseases related to mental healthcare in context with the elderly population. This is supported by the Alzheimer association report (2018) [7], which depicts the causes of

✉ Gagangeet Singh Aujla
gagangeet.aujla@newcastle.ac.uk;
gagi_aujla82@yahoo.com

¹ Chandigarh University, Mohali, India

² Newcastle University, Newcastle Upon Tyne, UK

³ Thapar Institute of Engineering and Technology, Patiala, India

death in the period of 2000–2015. Alzheimer disease becomes irreversible when it reaches later stages and the memory/thinking skills are destroyed that leads to difficulty in carrying out simple routine livelihood tasks for human being. The brain develops amyloid plaques and neurofibrillary tangles that were found after the brains of deceased were examined. According to the report, the reason of deaths has witnessed a growth of 123% for Alzheimer patients as compared to – 11% for heart disease and – 16% for stroke. Moreover, by 2050, the number of Alzheimer-affected patients is projected to surpass the mark of 15 million in USA alone, i.e., nearly two cases are expected to develop in a period of one minute [7]. Even more, according to the statistics released by National Center of Health Statistics (2017) [8], the number of deaths in USA which is attributed to the mental diseases like Alzheimer has shown an exponential growth. The increase in such deaths is attributed to the fact that the detection of Alzheimer in patients is often late due to which reversing the degradation becomes very difficult. Another reason is attributed to insufficient support mechanisms for such patients. These facts necessitate the need of reliable technological solutions for the detection and tracking (like visual tracking [9]) of the people suffering with Alzheimer's disease and thereafter providing necessary assistance to improve their quality of life.

2 Related work

In lieu of the above necessity, various existing proposals have tried to devise different solutions related to the prediction or detection of Alzheimer disease. The prediction of Alzheimer disease is being pursued via different approaches such as using MRI data and motion sensor data. For example, Chitradevi et al. [10] concentrated on brain subregions analysis for prediction of Alzheimer disease using optimization techniques. The research work carried out concluded that hippocampus biomarker acts as an important biomarker to analyze the Alzheimer disease. Alejandro et al. [11] used various deep learning techniques along with transfer learning to achieve better results on sagittal MRIs available in ADNI and OASIS datasets. The research initiatives undergoing for Alzheimer prediction are using either medical data (like MRI) or sensor data (like gait sensors, skeletal sensors). Long et al. [12] proposed a machine learning-based scheme to discriminate between Alzheimer patients and mild cognitive patients for healthy elderly cases.

In another work, Liu et al. [13] proposed a multi-template learning mechanism for the automated diagnosis and staging of Alzheimer's disease. For this purpose, a feature selection algorithm is used to model the relationships

between templates and patients followed by the usage of support vector machine classifier. Seifollahi et al. [14] performed experimental study on 24 elderly women. The experiment involved skeletal data analysis of 12 women who were affected with Alzheimer and 12 women who were healthy. The results showed that using support vector machines for feature extraction yielded good classification accuracy. The survey performed in [15] discusses various gait analysis approaches, pre-processing steps as well as various use cases where gait analysis can be explored further. In [16], the authors deployed a whole-brain hierarchical network to represent the patients based on automated anatomical labeling. They used feature selection algorithm to reduce the dimensionality and thereafter adopted multiple kernel boosting algorithm for the classification of the subjects. After analysis of these proposals, it can be concluded that machine learning algorithms have been used to detect and predict the Alzheimer's disease. However, the artificial intelligence-driven algorithm based on deep learning and cognitive intelligence can help to recognize or detect or predict the early symptoms of Alzheimer's disease in an efficient manner. Even more, the Alzheimer patients can be provided with an Internet of things (IoT)-based assistance for an enhanced quality of life.

In [17], a cloud- and IoT-based healthcare system has been proposed for predicting the serious diseases. However, the above proposal has not targeted the mental diseases like Alzheimer or dementia. Even more, the existing research has been limited to cloud-based systems and none of the proposals have tried to utilize the edge or fog devices for handling analytical tasks. Giampaolo et al. [18] collected behavioral data in a noninvasive way to achieve a machine learning classification which supports the vision of providing better healthcare services without affecting the lifestyle of the patients under monitoring. Research being undertaken in these areas supports the applicability of the model proposed in this research work in real life. However, none of the above-discussed existing proposals have realized the multi-varied solution for the detection, tracking and assistance of Alzheimer patients.

2.1 Contributions

To overcome these limitations, we proposed a deep learning-based Internet of Health Framework, called DeTrAs, for assisting Alzheimer patients. The contributions are listed as follows:

- A layered architecture for Internet of Health (IoH) ecosystem comprises fog data analytics. In this mode, a trigger-based sensor activation model is proposed for managing master and slave sensors.

- A recurrent neural network (RNN)-based Alzheimer prediction scheme is proposed which uses sensory movement data collected in IoH ecosystem.
- An ensemble approach for abnormality tracking for Alzheimer patients is designed which comprises two parts: (1) convolutional neural network (CNN)-based emotion detection scheme and (2) timestamp window-based natural language processing scheme.
- An IoT-based assistance mechanism is suggested for Alzheimer patients.

3 System model

The system model comprises a layered architecture for IoH system. The architecture is inspired from the fact that decision making in IoT environment [19] can be supported by machine learning techniques. Figure 1 depicts the layered architecture which consists of five layers described as follows:

User layer This layer consists of i users comprising human beings (Alzheimer and non-Alzheimer) at different locations in a smart environment.

IoT/Sensor layer This layer consists of j types of IoT sensors which are used to sense the data generated at user layer. These sensors can be embedded within the human body, inside a smart home and within the smart city infrastructure. The various types of sensor deployed in the proposed model are zenith sensor, voice recognition sensors, emotion sensor, acceleration sensor and medical sensors. These sensors capture heterogeneous data, i.e., auditory, visual, motion and medical.

Fog layer The data sensed by the sensors are forwarded to the computational layer, i.e., fog layer. This layer consists of k fog nodes or servers (equipped with sufficient computing resources or power) deployed at different geographical locations across a smart city. The entire computational or analytical tasks like data processing and analysis are performed closer to the location of the users, i.e., at fog layer.

Core network layer This layer comprises core network infrastructure, i.e., forwarding devices which are responsible for the entire transmission of data from fog layer to the cloud layer.

Cloud layer These sensed data are finally forwarded to cloud-based storage for future analysis. Moreover, if in a case, the fog devices do not have sufficient computing power to process the data, then such computation is passed to the cloud servers for processing. This layer compliments the resource constrained fog devices deployed in the proposed model.

3.1 Trigger-based sensor activation model

The deployed sensors have been divided into two categories, namely master (α) and slave (β) sensors. A trigger-based activation mechanism is used for handling these sensors in order to achieve improved energy efficiency. Initially, the master sensors are active, but the slave sensors are activated only as and when required. The slave sensors stay in a sleep mode due to non-active message received from the master sensor. For this purpose, the Message Queue Telemetry Transport (MQTT) publish/subscribe scheme is used to awake the slave sensors whenever required. Here, the patients are equipped with the gait sensors that are considered as master sensors to all other

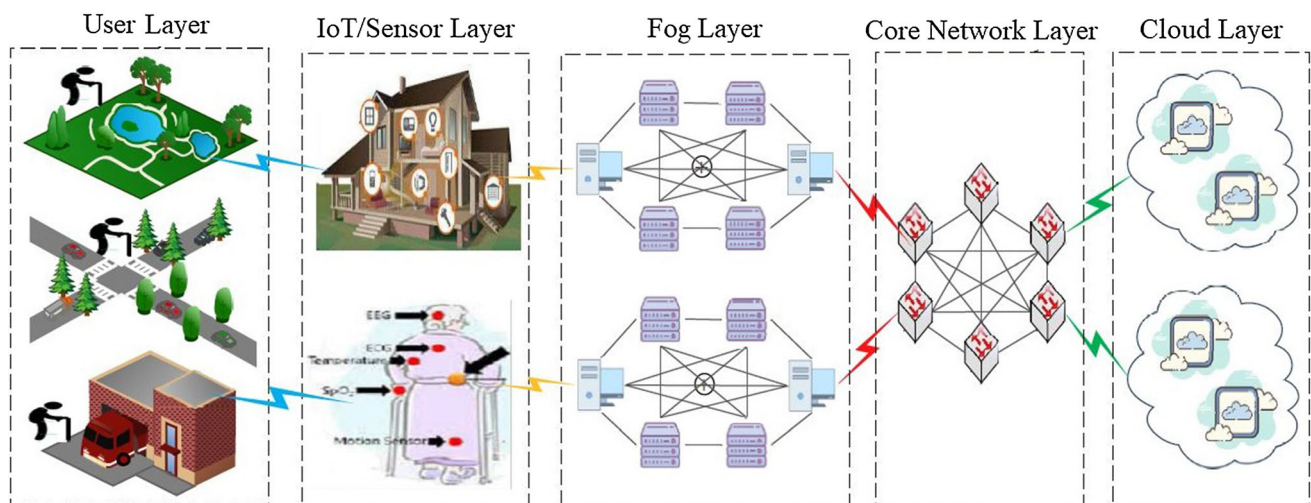


Fig. 1 Layered architecture for IoT-based healthcare framework

sensors. The gait sensor tracks the movement of patient which can further trigger the slave sensors when the patient is in motion. The set of slave sensors are listed as follows:

$$Z_{IG}, Z_{EG}, Z_{R_1}, Z_{R_2}, Z_{R_3}, \dots, Z_{R_n} \tag{1}$$

$$A_{IG}, A_{EG}, A_{R_1}, A_{R_2}, A_{R_3}, \dots, A_{R_n} \tag{2}$$

$$BL_{IG}, BL_{EG}, BL_{R_1}, BL_{R_2}, BL_{R_3}, \dots, BL_{R_n} \tag{3}$$

$$B_{O_1}, B_{O_2}, B_{O_3}, \dots, B_{O_M}; M_H; M_T \tag{4}$$

where Z represents the zenith sensor, IG represents the internal sensor placement, EG represents the external sensor placement, R_j represents the room numbers (1 to n), A represents the auditory sensor, BL represents the Bluetooth sensor boards, B represents the binary sensor, O_j represents the openable device sensors (1 to n), M_H represents the medical sensor for heartbeat and M_T represents the medical sensor for body temperature.

The considered sensory system is scalable in nature as the number of sensors is directly coupled with the number of rooms and openable devices. The number of sensors could be increased or decreased as per the need of the Alzheimer patient surrounding environment. All the slave sensors mentioned above subscribe to the master sensor which publishes the message of activation or non-activation as per the requirements. The slave sensors on receiving the non-active message enter the sleep mode for energy saving and start sensing once they receive the activation trigger. The important aspect of MQTT mechanism is that the publisher is decoupled from the subscriber. The analysis of the data sensed by the slave sensors is performed in synchronization with the master sensors. The sensed data are generally in a raw form and require pre-processing for better efficiency and accuracy. The following function has been defined for the activation of sensors.

$$f_{\beta} = \begin{cases} 1 : & \text{If } \alpha \rightarrow \text{ACTIVATE} \\ 0 : & \text{Otherwise} \end{cases} \tag{5}$$

4 DeTrAs: proposed scheme

The proposed scheme, i.e., DeTrAs, is divided into three phases (as shown in Fig. 2) which are discussed as follows:

4.1 Phase 1: RNN-based Alzheimer prediction

The first phase involves the task to identify whether the subject under supervision is an Alzheimer patient or not. For this purpose, RNN, a popular deep learning approach, has been used. RNN comprises numerous processing layers with agenda to learn about representations of data using abstraction [20]. To predict the patients with Alzheimer,

the input data sensed by the sensory system are fed to the RNN model. The model for Alzheimer disease prediction is defined as follows [20]:

$$s^t = f(s^t - 1, r^t; \theta) \tag{6}$$

where $s(t)$ represents the current hidden layer which is a function of the previous hidden layer, $r(t)$ denotes the current input and θ represents the various parameters of the input functions.

Based on the above definition, the graphical representation of the RNN model is shown in Fig. 3.

The complete set of parameters considered for training the RNN model are given in Table 1.

RNN assigns different weights to the above-defined parameters in multiple hidden layers while performing supervised learning to train the model. The trained model is then installed on the computational fog devices which receives an input from the master sensors, i.e., gait sensors embedded on the subjects for the purpose of movement monitoring as shown in Fig. 2. The fog receivers depend on the location of the subject, i.e., the various areas where the subject could be found such as home environment, public areas, local businesses and public transportation. The sensed data are continuously transmitted to the fog devices till subject is identified. Jayneel et al. [22] proposed similar technique of patient monitoring using fog computational devices for assisting patients in their home environments. The results obtained showed better performance in terms of latency and data overloading for fog computing-based environment instead of non-fog computing-based environment. Inspired from the same, we have used the fog layer to enforce latency-conscious system.

A unique id (P_{id}) is assigned to the subject which is transmitted to the cloud via network gateway, and the profile type with the respective id is updated accordingly. If a patient is detected as positive, then the next phase of DeTrAs is triggered for behavior abnormality using multiple sensors.

The weights of RNN are optimized to achieve a better classification result by using supervised learning for adapting the weights, which in turn leads to dynamic behavior of RNN. The mean squared error is focused here for minimizing the classification error $E^{(class)}$.

$$E^{(MSE,c)} = \frac{1}{N \cdot T} \sum_{n=1}^N \sum_{t=1}^T \sum_{n=1}^{N^{(c)}} \left(x_{nti}^{(c)} - y_{nti}^{(c)} \right)^2 \tag{7}$$

The number of time series in the training dataset is represented by N , and T is the number of time steps in every time series. The evaluation involved taking a time step of 50 s for being able to capture long-term dependencies in data and is represented by t . The feature vectors taken into consideration are 9 and are presented by αL . The variables

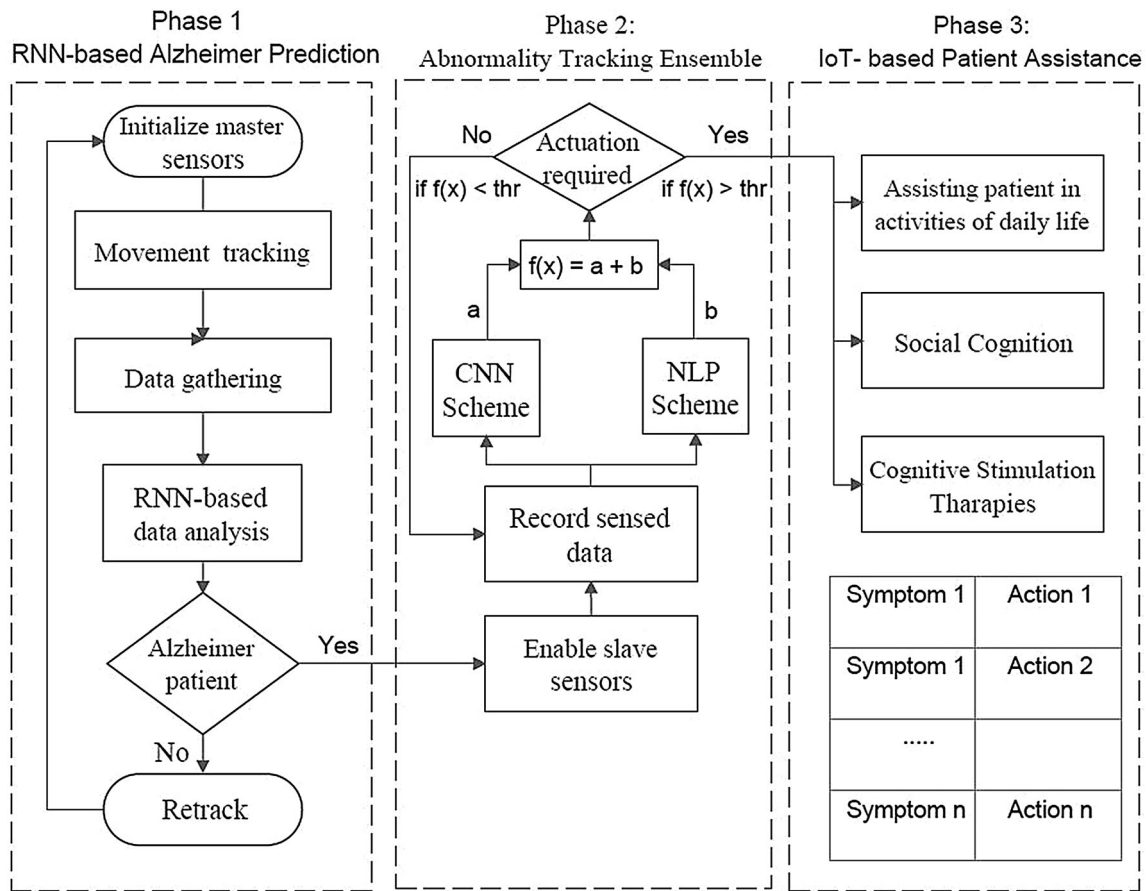


Fig. 2 DeTrAs framework

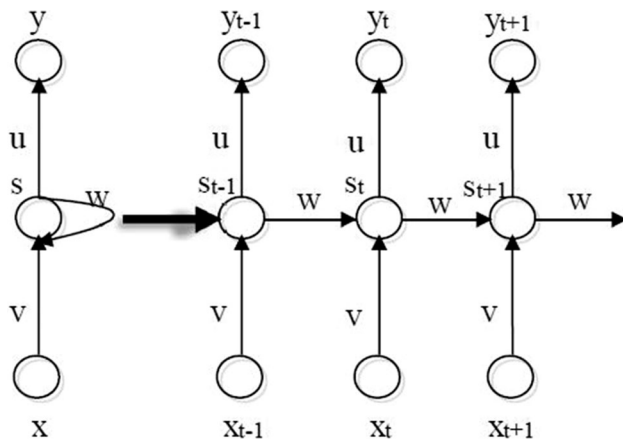


Fig. 3 RNN model

$y_{nti}^{(c)}$ and $x_{nti}^{(c)}$ are defined as the output and target value of the i th classification neuron at the t th time step of the n th time series. The number of neurons taken is 50 and is represented using NU.

The proposed framework is focused on continuous monitoring of the candidate in an IoT-enabled environment till the patient with Alzheimer is detected. The candidate

Table 1 Notation table [21]

Notation	Description
AA _{HF}	Ankle acceleration—horizontal forward
AA _V	Ankle acceleration—vertical
AA _{HL}	Ankle acceleration—horizontal lateral
UL _{HF}	Upper leg—horizontal forward
UL _V	Upper leg—vertical
UL _{HL}	Upper leg—horizontal lateral
TA _{HF}	Truck acceleration—horizontal forward
TA _V	Truck acceleration—vertical
TA _{HL}	Truck acceleration—horizontal lateral

continues to perform regular activities without any interference. The movement tracking of the patient is done via various means, i.e., gait sensors as well as Bluetooth board sensors. Tracking of the candidate's location in the house is performed by the Bluetooth board sensors [23], whereas gait sensors track the movement style of the candidate. The work flow presented in Algorithm 1 is controlled through a decision variable (μ). This variable depicts a positive

Alzheimer case if its value is 1 and negative case if its value is 0 represented as follows:

$$\mu = \begin{cases} 1 : \text{Affected} \\ 0 : \text{Not-affected} \end{cases} \quad (8)$$

Algorithm 1 RNN-based Alzheimer prediction

```

Input:  $\alpha$ 
Output:  $\mu$ 
Initialize  $\mu \leftarrow 0$ 
1: while  $\mu \neq 1$  do
2:   Declare  $input[T][\alpha L]$ 
3:   for  $t=1 : T$  do
4:     Declare  $s[\alpha L]$ 
5:     for  $\alpha$  sensors do
6:        $s[\alpha L] \leftarrow \alpha$ 
7:     end for
8:      $input[t] \leftarrow s$ 
9:   end for
10:  Declare  $neurons[T]$ 
11:  for  $NU = 1 : T$  do
12:     $neurons[NU] \leftarrow input[NU]$ 
13:    Apply hyperbolic tangent activation function
14:    Use a dropout of 0.5 to reduce over-fitting
15:  end for
16:  Output Class = LSTM-RNN Model Output
17:  if (Output Class == Freeze) then
18:     $\mu \leftarrow 1$ 
19:  end if
20: end while

```

In this algorithm, the input from gait sensors is collected in batch of 50 s represented as follows:

$$\alpha = \left(AA_{HF(1,2,\dots,n)}, AA_{V(1,2,\dots,n)}, AA_{L(1,2,\dots,n)}, \right. \\ UL_{HF(1,2,\dots,n)}, UL_{V(1,2,\dots,n)}, UL_{L(1,2,\dots,n)}, \\ \left. TA_{HF(1,2,\dots,n)}, TA_{V(1,2,\dots,n)}, TA_{L(1,2,\dots,n)} \right) \quad (9)$$

The data are batched with an objective of analyzing the temporal trends, and hence, LSTM RNN is used where data are passed on to 50 declared neurons of the RNN. The activation function used in the process is hyperbolic tangent, and a dropout of 0.5 is taken to avoid over-fitting. The output of this algorithm for a class can be either *Freeze* or *Normal* values. Once the output is received, the control is passed on to phase 2 if $\mu = 1$ or the loop is repeated again.

4.2 Phase 2: ensemble approach for abnormality tracking

In this phase, an ensemble approach for tracking of Alzheimer patient abnormality is designed. This is done to reduce the number of incorrect detection made by the proposed scheme as the weight assigned to different input sources of detection will be eliminating the outliers. This phase considered two cases: (1) video (human pose

analysis [24], person and visual identification [25–27], action recognition [28]) and (2) audio data collected from sensors (linguistic analysis [29, 30]). The output of this phase is obtained on the basis of the following function:

$$f(x) = a + b \quad (10)$$

where $a = W * SA$ and $b = W * SB$ and SA is output of case 1 and SB is output of case 2.

In the proposed scheme, the weights assigned for each of the cases are given as follows:

$$f(W) = \begin{cases} W = 0.6, & \text{If sensor type = Audio} \\ W = 0.4, & \text{If sensor type = Video} \end{cases} \quad (11)$$

If the combined weight of the sensory alarm generation is more than the prescribed threshold, the alarm is considered to be a true alarm and hence the mechanisms would be triggered for assisting the Alzheimer patient. The proposed schemes for both these cases are described as follows:

4.2.1 Case 1: CNN-based emotion detection

The composite video data from zenith sensors are converted to time-variant images of 120 * 160 RGB images. These images would then fed to the CNN model to detect the emotion of the subject [31]. The major emotion categories and their relative priorities defined in [32] are shown in Table 2.

The proposed scheme uses a pre-trained CNN model as shown in Fig. 4 for extracting the features of the dataset. The extracted representation is fed as an input to the CNN classifier which analyzes the time-variant images fetched from zenith sensors. CNN model comprises an input layer, numerous hidden layers and an output layer. Hidden layers are further bifurcated into pooling layers, convolutional layers and a completely connected layer. The kernel (mask) is applied on a part of input vector (receptive field) in the convolution layer. On application of the kernel, a weighted value of a part of the input vector is calculated and is fed into the hidden layer. This process continues incrementally to cover all the parts of the input vector. Hence, there is a hidden neuron present for each receptive field in the hidden layer which does the task of learning the overall bias in the hidden layer. The mathematical representation of the

Table 2 Emotion categories

Emotion	Priority
Happiness	Normal
Sadness	Normal
Disgust	Normal
Surprise	High
Fear	Very high
Anger	Very high

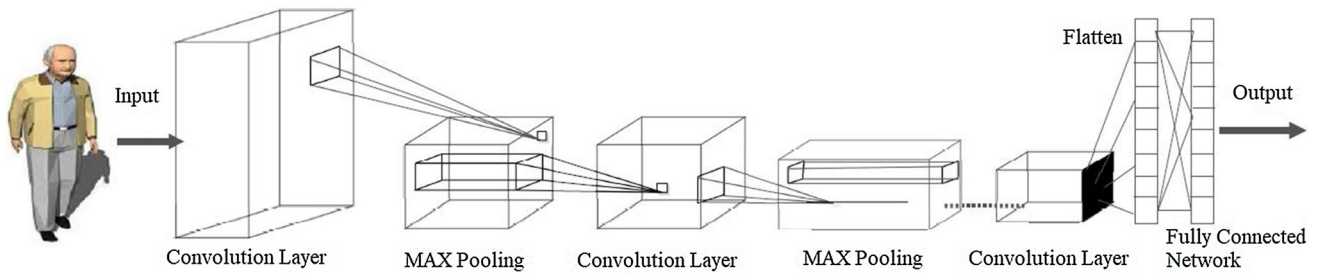


Fig. 4 Tracking of Alzheimer patient

output of hidden neuron post-convolution can be represented as follows [33, 34]:

$$\text{conv}(w, \alpha)_{xy} = \sigma \left(b + \sum_{i=0}^n \sum_{j=0}^n w_{(i,j)} \alpha_{(x+i,y+j)} \right) \quad (12)$$

where σ is the activation function, b is the value of bias, w is the weighted mask of dimension $n \times n$ and $\alpha_{(x,y)}$ represents the input activation at (x, y) th position.

On activation of all the slave sensors, the zenith sensors start recording the Alzheimer patient facial features data. The subject under observation may express emotion of surprise and confusion if he/she is suffering from effects of Alzheimer. On detection of such an emotion, an alarm is generated for triggering actuation. This phase generates an analytical result, SA. To avoid false alarms, the audio data are also analyzed for abnormality tracking in order to reach a firm conclusion.

4.2.2 Case 2: Timestamp window-based natural language processing scheme

In this case, the data collected by auditory sensors are used to validate the partial alarm generated in the previous case on the basis of zenith sensors. The timestamp window mechanism is used to filter out the audio stream fetched from the auditory sensors. A window of 15 s is proposed from the timestamp where video frame recognizes a triggering emotion. The speech-to-text synthesizer is used to convert the given speech to the textual form. Then, the pre-processing steps (tokenization, stop words removal, lemmatization) are performed on the textual data. Parts-of-speech (POS) tagger is used to further analyze the structure of the sentences (like sentiment analysis [35]). The existing corpus is fed to the naive Bayes classifier for the identification of sentences which represents the questions triggered to extract the information about identity of the visitor from the cloud repository. The analytical resultant of this phase is represented as SB. The overall flow of this case is depicted in Fig. 5.

The location of the user is tracked using the wearable smart band which provides GPS coordinates. The video

data received from zenith sensor are converted into still image frames which would be represented in the form of $m * n$ and audio data generated from audio sensor for every 15 s converted in the form of text. The current location of the user and time for which the user has stayed in that room is used to activate the various sensors attached in that room.

The inputs from activated zenith and audio sensors are analyzed in order to evaluate the need of actuation. This ensemble approach for this entire task is represented using Algorithm 2. Here, a decision variable (γ) is used to control the actuation represented as follows:

$$\gamma = \left\{ \begin{array}{l} 1 : \text{Actuation required} \\ 0 : \text{Actuation not-required} \end{array} \right\}. \quad (13)$$

Algorithm 2 is designed to depict the workflow of the ensemble approach. The algorithm involves conversion of video-based data (represented as α) into image frames which are then fed to the pooling layer.

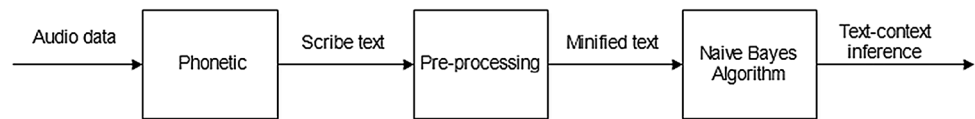
$$\alpha = (Z_{IG}, Z_{EG}, Z_{R_{(1,2,\dots,n)}}). \quad (14)$$

Similarly, the data from audio sensors (represented as β) are fed to the synthesizer which generates respective text-based data.

$$\beta = (A_{(1,2,3,\dots,n)}). \quad (15)$$

On image frames, the pooling layer reduces the data over selected feature vectors. The CNN model with sigmoid activation function generates the output class representing emotion of the object in frame. Similarly, the text pre-processing techniques such as stemming/lemmatization, tokenization and stop word removal are applied to convert the data into high-priority tokens only. Then, naive Bayes classifier is used to evaluate the emotion. If both zenith and audio sensors are in consensus with the emotion detected, then the threshold value is checked in order to reduce the probability of false alarms. Depending upon the type of emotion and the location of user, the required assistive activity is launched as discussed in phase 3.

Fig. 5 NLP-based sentiment analysis



Algorithm 2 Abnormality tracking of Alzheimer patients

Input: α, β

Output: γ

```

1: Declare  $\alpha, \beta$ 
2: while  $\gamma \neq 1$  do
3:   for Zenith sensors do
4:      $[I_{m*n}] \leftarrow \alpha$ 
5:   end for
6:   for Audio sensors do
7:      $[J_m] \leftarrow \beta$ 
8:   end for
9:   Pooling Layer:  $P_{p*q} \leftarrow [I_{m*n}]$ 
10:  Apply Sigmoid activation function on  $[I_{m*n}]$ 
11:  Output Class = CNN Classifier Model Output
12:   $z_1 =$  Output Class
13:  Declare input as  $[J_m]$ 
14:  Audio data  $J_m$  converted into Text  $[T_m]$ 
15:  Apply Naive-Bayes Classifier on  $[T_m]$ 
16:  Output Class = Naive-Bayes Model Output
17:   $z_2 =$  Output Class
18:  Calculate  $\gamma = (w_1 * z_1 + w_2 * z_2)$ 
19:  if  $(\gamma == 1)$  then
20:    Go to Phase 3
21:  end if
22: end while
  
```

4.3 Phase 3: IoT-based patient assistance mechanism

The third phase of the scheme involves the usage of IoT devices for providing different types of assistance to the Alzheimer patient when a trigger is generated by phase 2 (such as social community message propagation [36]). The assistance could be of various types listed as follows:

- In corrective assistance, the Alzheimer patient is informed about the decision which is incorrect and needs a corrective action or should be reverted.
- In reinforcing assistance, the process opted for an incorrectly performed action in the past is reiterated so as to perform it correctly at the current instance. Reinforcing assistance is aimed at cognitive stimulation for delaying the effects of Alzheimer's disease.
- Supportive assistance involves the case when requirement is triggered by Alzheimer's patient using an application.

In lieu of the above, the IoT-based assistance can be classified on the basis of complexity of trigger and support mechanisms into activities of daily life (ADL), social cognition and cognitive stimulation therapy. The different possible scenarios for each of the above-mentioned categories are discussed as follows:

4.3.1 ADL

These are routine activities which are performed by all elderly people and mostly include Alzheimer patient as an individual. The different types of ADL are discussed as follows:

- Meal tracking activity is to ensure that the patient does not forget to take any of the meals in a day. For example, the motion sensors embedded in a wearable band enable the detection using the arm motions to know whether meal eating motions have been performed or not.
- Bath activity is embedded to ensure that the patient has taken bath in the day. Binary sensors embedded on the bathroom utilities and house gate help to detect whether bath activity has been performed or not. In a case the Alzheimer patient has forgotten to take bath and tries to open gate of the house, a reminder would be sent for the completion of bath activity before leaving the house.
- Medicine activity is used to ensure that the patient has taken the prescribed medicines. Binary sensor embedded on the medicine box detects whether the pills have been used or not. If this activity is skipped, then an initial reminder is sent to the patient and thereafter a second reminder is sent to the family members of the patient.
- Hydration activity ensures that the patient is consuming the water in an appropriate amount. The motion sensors embedded on the wearable band detect the number of times the motion-based activity for drinking water has been performed. If in a case, the patient has not drunk the desirable amount of water to keep him hydrated, then a trigger is generated.
- Safe area fencing ensures that an Alzheimer patient does not enter any area marked as unsafe. GPS sensor data are used to match the unsafe areas coordinates with the current location of a patient. If in a case the patient enters an unsafe area, a trigger is sent to the family members via cloud layer.

4.3.2 Social cognition

These activities are performed when an Alzheimer patient is interacting with other human beings, i.e., a relative, family member or unknown person. This is comparatively complex activity as it involves the tasks with relatively

larger computational requirements. The different types of social cognition activities are listed as follows:

- *Visitor recognition* is an activity where the Alzheimer patient requires assistance to recognize the visitors. If the ZIG and AIG sensors generate the trigger of inability of patient to recognize the visitor, then the ZEG sensor data are used to check whether the visitor belongs to the patient known list of humans or not. Accordingly, the patient is presented with the related information about the visitor on the smart glasses so that he can handle the situation.
- *Task scheduler activity*: The A_{IG} , A_{EG} , A_{R1} , A_{R2} , A_{R3}, \dots, A_{RN} sensors are used to continuously process the audio stream. The NLP process discussed in Section 4.2.2 is applied to find any time-based events scheduled for the future. On detection of any such event, the information is passed to the cloud layer which sends a reminder on the basis of the stored database.
- *Conversation recaller*: The A_{IG} , A_{EG} , A_{R1} , A_{R2} , A_{R3}, \dots, A_{RN} sensors are used to find any reference to the past event. In case, if there is a reference to any past event and the patient is not showing any sign of recall, then the existing conversations related to the visitor with the same context are shared with the patient on audio speaker.

4.3.3 Cognitive stimulation therapy

These activities aim at reducing the rate of impact of Alzheimer disease on the patient. This IoT-based support mechanism is triggered on the basis of the inputs from wearable and zenith sensors. Wearable sensors are used to judge whether the patient is awake or not and the zenith sensors are used to conclude whether the patient is free or not. If in a case the patient is awake and free, then the cognitive stimulation games can be presented to the patient via the home theater system or smart phone.

In this way, the IoT-based assistance mechanism is used to improve the quality life index of the Alzheimer patient.

The flow of IoT-based patient assistance mechanism is depicted in Fig. 6.

Also, Table 3 describes the different verticals and cases of the suggested IoT-based patient assistance mechanism.

5 Performance evaluation

DeTrAs involves the amalgamation of different existing techniques to provide an ecosystem to the Alzheimer’s patients which can assist them in day-to-day activities. Different datasets for different phases have been used for evaluating the effectiveness of DeTrAs in a simulated environment.

In phase 1, the Daphnet dataset [21] has been used to predict the Alzheimer’s disease. A total of 1,025,891 records were used to train the RNN model along with various classification algorithms (Bayes Net, naive Bayes, logistic, simple logistic, decision tree). The different evaluation results are depicted in Table 4 in terms of correctly classified instances, true-positive rate, false-positive rate, precision, recall and F-score. The superiority of RNN in terms of precision, recall and F-score is clearly evident from Fig. 7. Daphnet dataset is devised to benchmark the automatic detection mechanisms for Freezing of Gait. The correct classification instances percentage of RNN (DeTrAs) is 88.59%, which is 6% higher than nearest candidate, i.e., decision tree as well as the precision of RNN (DeTrAs) is higher than decision tree and Bayes Net on Daphnet dataset.

Even more, the accuracy of RNN classifier is shown in Table 5. RNN classifier is found to be performing better than decision tree achieving 88.63 percent accuracy. RNN classifier performs better for this case because in case any exceptional data values are there in the data then the decision trees fall apart which leads to reconstruction of tree, whereas the same can be achieved in RNN by adjusting the weight factors.

In phase 2, the ensemble approach is used for emotion extraction using the video and audio data feeds. For video feed, the MMI dataset [37] is used as reference for

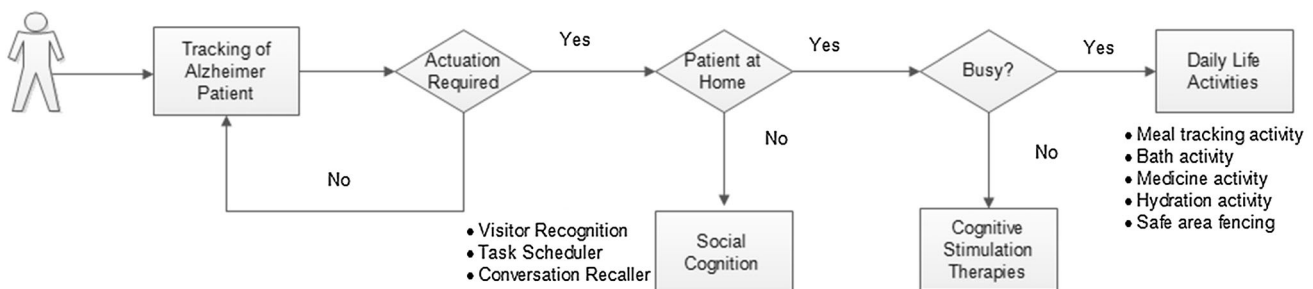


Fig. 6 Flow of IoT-based patient assistance mechanism

Table 3 IoT-based patient assistance mechanism

Name of activity	Triggered sensor	Actuation	Computational layer	Type of activity
Meals	Motion	Wearable band reminder	Edge layer	ADL
Bath	Binary	Audio speaker	Edge layer	ADL
Medicine	Binary	Band reminder, reminder to family members	Edge/cloud layer	ADL
Hydration	Motion	Wearable band reminder	Edge/cloud layer	ADL
Safe area fencing	GPS	Band reminder, reminder to family members	Cloud layer	ADL
Visitor recognition	Zenith and audio	Smart glass recognizer	Edge/cloud layer	Social cognition
Task scheduler	Audio	Wearable band reminder	Edge/cloud layer	Social cognition
Conversation recaller	Audio	Audio speaker	Edge/cloud layer	Social cognition
Cognitive stimulation therapies	Wearable and Zenith	Cognitive games	Edge/cloud layer	Cognitive stimulation

Table 4 Alzheimer detection using different algorithms

Algorithm	CCI (%)	TP rate	FP rate	Precision	Recall	F-score
Bayes Net	72.8	0.729	0.203	0.8	0.729	0.742
Naive Bayes	72.26	0.723	0.267	0.751	0.723	0.709
Logistic	74.16	0.742	0.276	0.753	0.742	0.71
Simple logistic	74.32	0.743	0.281	0.748	0.743	0.708
Decision tree	82.88	0.829	0.164	0.807	0.829	0.81
RNN (DeTrAs)	88.59	0.886	0.866	0.83	0.886	0.84

Fig. 7 RNN versus other classification algorithms

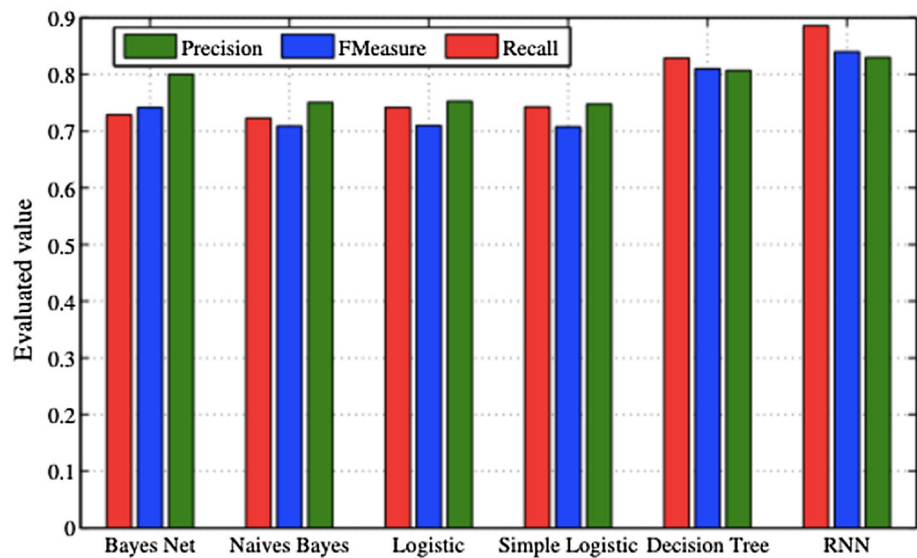


Table 5 Freezing of gait classification accuracy

Name of classifier	Accuracy
Decision tree	82.68
RNN (DeTrAs)	88.63

extracting the emotion from the image data. The proposed CNN-based tracking is compared with popular SVM classifier. CNN is found to be more accurate in contrast to SVM in case of emotion detection from image feed as shown in Table 6. The CNN performs better than SVM because of the fact that the CNN performs better at finding

Table 6 Phase 2 video feed evaluation results

Name of classifier	Accuracy
SVM	62.89
Deeper CNN (DeTrAs)	79.9

local connections with the help of available filters as well as the time taken in testing of CNN is much smaller as compared to SVM, whereas the same is opposite for training phase.

For extensive evaluation, the CNN model (DeTrAs) is pre-trained on the FER dataset [38]. The FER dataset is divided into training (28,700 images), testing (3500 images) and validating (3500 images) phases. The initial network learning rate and minimum learning rates are set to 0.02 and 0.001, respectively. Here, we used stochastic pooling which leads to random sample selection from the training set in each of the training epochs. It was observed that network does not overfit due to the selection of stochastic pooling dropout. The epoch with best training accuracy has been selected as our trained model. Table 7 shows the variation of accuracy with an increase in the number of epochs for training, testing and validation sets.

Table 8 shows the comparative analysis of classification accuracy for FER dataset using three approaches: (1) six randomly initialized single (RIS) CNN models, (2) average ensemble of these six networks and (3) hinge loss framework. It depicts the variations in the classification accuracy for validation and testing based on different approaches.

Similarly, the comparative analysis of classification accuracy three approaches for SFEW dataset [39] is shown in Table 9. Here also the variations in terms of classification accuracy for validation and testing have been compared based on different approaches.

Figure 8 depicts the comparative representation of classification accuracy for these three models for FER and SFEW datasets.

For audio feed, the sensors are first synthesized to scribe text which is then pre-processed before initiating the classification process. The dataset provided in [40] is used to evaluate the performance of the emotion detection from text data. Two classes of classification algorithms are

Table 7 Training and testing accuracies with an increase in epochs on FER

No. of epochs	Training accuracy (%)	Non-voting accuracy (%)	Voting accuracy (%)
10	40.78	40.13	40.56
50	74	64.2	66.08
100	78.45	65.26	68.04
150	86.67	70.84	75.26
180	90.22	74.46	81.04

Table 8 Classification accuracy of different models on FER dataset

Model	Six RIS CNN models	Average ensemble of six RIS CNN models	Hinge loss CNN framework
Validation	68.72	73.24	77.92
Test	70.45	76.37	79.86

Table 9 Classification accuracy of different models on SFEW dataset

Model	Six randomly initialized single (RIS) CNN models	Average ensemble of six RIS CNN models	Hinge loss CNN framework
Validation	53.46	58.35	66.58
Test	60.04	63.47	71.02

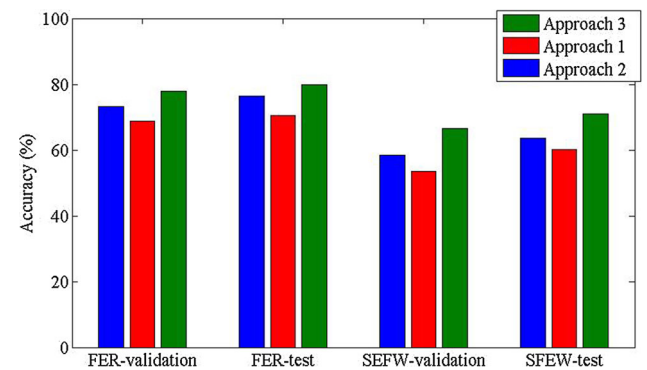


Fig. 8 Evaluation results for CNN (DeTrAs)

Table 10 Phase 2 audio feed evaluation results

Name of classifier	Accuracy
Decision tree	71.83
Naive Bayes (DeTrAs)	75.02

evaluated in the experiment, i.e., naive Bayes and decision tree. Naive Bayes classifier is found to be more accurate in comparison with the other variants as shown in Table 10.

The emotion detection using two modes of data, i.e., video and audio feed, helps to reinforce the correct detection of assistance situations and hence reduces the number of false alarms. After the analysis of the evaluation results, it is quite evident that DeTrAs acts an accurate framework for providing correct classification Alzheimer patients which are provided with assistive care using IoT-based mechanism.

6 Conclusion and future scope

This paper presented an ensemble framework DeTrAs: deep learning mechanism for predictive and assistive care for Alzheimer patients in Internet of Health ecosystem. In DeTrAs, three different phases work in tandem with each other. The first phase predicts the Alzheimer patients on the basis of emotional moments of the subject collected using sensors deployed. In phase 2, the predicted subjects are further evaluated for the abnormality staging on the basis of video and audio feeds with the larger goal of providing assistive healthcare provisions in the last phase. DeTrAs has been evaluated using different datasets on the basis of accuracy, precision, recall, true-positive rate, false-positive rate, F-score and correctly classified instances. DeTrAs outperforms the existing variants of its category due to deeper training using multiple layers of neural networks. The proposed approach could be further enhanced using ambient intelligence and game theoretic approaches to achieve Nash equilibrium and thus increase the performance of DeTrAs as well.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kumar N, Chilamkurti N, Misra SC (2015) Bayesian coalition game for the Internet of things: an ambient intelligence-based evaluation. *IEEE Commun Mag* 53(1):48–55
- Aujla GS, Chaudhary R, Kaur K, Garg S, Kumar N, Ranjan R (2018) Safe: Sdn-assisted framework for edge-cloud interplay in secure healthcare ecosystem. *IEEE Trans Ind Inform* 15(1):469–480
- Sharma D, Aujla G, Singh, Bajaj R (2019) Evolution from ancient medication to human-centered healthcare 4.0: a review on health care recommender systems. *Int J Commun Syst*, e4058
- Aujla GS, Jindal A, Chaudhary R, Kumar N, Vashist S, Sharma N, Obaidat MS (2019) DLrs: deep learning-based recommender system for smart healthcare ecosystem. In: *ICC 2019-2019 IEEE international conference on communications (ICC)*. IEEE, pp 1–6
- Chaudhary R, Jindal A, Aujla GS, Kumar N, Das AK, Saxena N (2018) Lscsh: lattice-based secure cryptosystem for smart healthcare in smart cities environment. *IEEE Commun Mag* 56(4):24–32
- Hollis C, Morriss R, Martin J, Amani S, Cotton R, Denis M, Lewis S (2015) Technological innovations in mental healthcare: harnessing the digital revolution. *Br J Psychiatry* 206(4):263–265
- Association A et al (2018) 2018 Alzheimer's disease facts and figures. *Alzheimer's Dement* 14(3):367–429
- Heron M (2007) National vital statistics reports. National Center for Health Statistics
- Chen Y, Wang J, Xia R, Zhang Q, Cao Z, Yang K (2019) The visual object tracking algorithm research based on adaptive combination kernel. *J Ambient Intell Humaniz Comput* 10(12):4855–4867
- Chitradevi D, Prabha S (2020) Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease. *Appl Soft Comput* 86:105857
- Castro AP, Fernandez-Blanco E, Pazos A, Munteanu CR (2020) Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Comput Biol Med* 120:103764
- Long X, Chen L, Jiang C, Zhang L, Initiative ADN et al (2017) Prediction and classification of Alzheimer disease based on quantification of MRI deformation. *PLoS ONE* 12(3):e0173372
- Liu M, Zhang D, Shen D (2016) Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment. *IEEE Trans Med Imaging* 35(6):1463–1474
- Seifollahi M, Soltanizadeh H, Hassani Mehraban A, Fa Khamseh (2019) Detection of Alzheimer's disease in elder people using gait analysis and kinect camera. *J Health Biomed Inform* 6(3):178–196
- Prakash C, Kumar R, Mittal N (2018) Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artif Intell Rev* 49(1):1–40
- Liu J, Li M, Lan W, Wu F-X, Pan Y, Wang J (2016) Classification of Alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Trans Comput Biol Bioinform* 15(2):624–632
- Kumar PM, Lokesh S, Varatharajan R, Babu GC, Parthasarathy P (2018) Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier. *Future Gener Comput Syst* 86:527–534
- Casolla G, Cuomo S, Cola VSd, Piccialli F (2020) Exploring unsupervised learning techniques for the Internet of things. *IEEE Trans Ind Inform* 16(4):2621–2628

19. Piccialli F, Casolla G, Cuomo S, Giampaolo F, di Cola VS (2020) Decision making in IoT environment through unsupervised learning. *IEEE Intell Syst* 35(1):27–35
20. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436
21. Bächlin M, Plotnik M, Roggen D, Maidan I, Hausdorff JM, Giladi N, Tröster G (2010) Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. *IEEE Trans Inf Technol Biomed* 14(2):436–446
22. Vora J, Tanwar S, Tyagi S, Kumar N, Rodrigues JJ (2017) Faal: fog computing-based patient monitoring system for ambient assisted living. In: *IEEE 19th international conference on e-health networking, applications and services (Healthcom)*, vol 2017. IEEE, pp 1–6
23. Piccialli F, Cuomo S, di Cola VS, Casolla G (2019) A machine learning approach for IoT cultural data. *J Ambient Intell Humaniz Comput*, 1–12
24. Li F, Zhou S-R, Zhang J-M, Zhang D-Y, Xiang L-Y (2013) Attribute-based knowledge transfer learning for human pose estimation. *Neurocomputing* 116:301–310
25. Zhou S, Ke M, Luo P (2019) Multi-camera transfer GAN for person re-identification. *J Vis Commun Image Represent* 59:393–400
26. Gui Y, Zeng G (2020) Joint learning of visual and spatial features for edit propagation from a single image. *Vis Comput* 36(3):469–482
27. Long M, Peng F, Zhu Y (2019) Identifying natural images and computer generated graphics based on binary similarity measures of PRNU. *Multimed Tools Appl* 78(1):489–506
28. Zhou S, Chen L, Sugumaran V (2020) Hidden two-stream collaborative learning network for action recognition. *CMC Comput Mater Contin* 63(3):1545–1561
29. Xiang L, Wang X, Yang C, Liu P (2017) A novel linguistic steganography based on synonym run-length encoding. *IEICE Trans Inf Syst* 100(2):313–322
30. Xiang L, Wu W, Li X, Yang C (2018) A linguistic steganography based on word indexing compression and candidate selection. *Multimed Tools Appl* 77(21):28-969–28-989
31. Zhang D, Li Q, Yang G, Li L, Sun X (2017) Detection of image seam carving by using weber local descriptor and local binary patterns. *J Inf Secur Appl* 36:135–144
32. Edman P, Friesen WV, Ellsworth P (1972) *Emotion in the human face: guidelines for research and an integration of findings*. Elsevier, Amsterdam
33. Jindal A, Aujla GS, Kumar N, Prodan R, Obaidat MS (2018) Drums: demand response management in a smart city using deep learning and SVR. In: *IEEE global communications conference (GLOBECOM)*, vol 2018. IEEE, pp 1–6
34. Gulati A, Aujla GS, Chaudhary R, Kumar N, Obaidat MS (2018) Deep learning-based content centric data dissemination scheme for internet of vehicles. In: *2018 IEEE international conference on communications (ICC)*. IEEE, pp 1–6
35. Zeng D, Dai Y, Li F, Wang J, Sangaiah AK (2019) Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism. *J Intell Fuzzy Syst* 36(5):3971–3980
36. Gu K, Wang L, Yin B (2019) Social community detection and message propagation scheme based on personal willingness in social network. *Soft Comput* 23(15):6267–6285
37. Mollahosseini A, Chan D, Mahoor MH (2016) Going deeper in facial expression recognition using deep neural networks. In: *IEEE winter conference on applications of computer vision (WACV)*, vol 2016. IEEE, pp 1–10
38. Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B, Cukierski W, Tang Y, Thaler D, Lee D-H et al (2013) Challenges in representation learning: a report on three machine learning contests. In: *International conference on neural information processing*. Springer, pp 117–124
39. Dhall A, Goecke R, Lucey S, Gedeon T (2011) Static facial expression analysis in tough conditions: data, evaluation protocol and benchmark. In: *2011 IEEE international conference on computer vision workshops (ICCV workshops)*. IEEE, pp 2106–2112
40. Chaffar S, Inkpen D (2011) Using a heterogeneous dataset for emotion analysis in text. In: *Canadian conference on artificial intelligence*. Springer, pp 62–67

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.