IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 19, 2021, accepted June 27, 2021, date of publication July 2, 2021, date of current version August 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3094243

Internet of Things and Deep Learning Enabled Elderly Fall Detection Model for Smart Homecare

THAVAVEL VAIYAPURI[1], (Member, IEEE), E. LAXMI LYDIA[2], MOHAMED YACIN SIKKANDAR[3], VICENTE GARCÍA DÍAZ[4], IRINA V. PUSTOKHINA[5], AND DENIS A. PUSTOKHIN[6]

[1]College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 16278, Saudi Arabia
[2]Department of Computer Science and Engineering, Vignan's Institute of Information Technology (Autonomous), Visakhapatnam 530049, India
[3]Department of Medical Equipment Technology, College of Applied Medical Sciences, Majmaah University, Al Majma'ah 11952, Saudi Arabia
[4]Department of Computer Science, School of Computer Science Engineering, University of Oviedo, 33003 Oviedo, Spain
[5]Department of Entrepreneurship and Logistics, Plekhanov Russian University of Economics, 117997 Moscow, Russia
[6]Department of Logistics, State University of Management, 109542 Moscow, Russia

Corresponding author: Thavavel Vaiyapuri (t.thangam@psau.edu.sa)

ABSTRACT Recently, the techniques of Internet of Things (IoT) and mobile communications have been developed to gather human and environment information data for a variety of intelligent services and applications. Remote monitoring of elderly and disabled people living in smart homes is highly challenging due to probable accidents which might occur due to daily activities such as falls. For elderly people, fall is considered as a major reason for death of post-traumatic complication. So, early identification of elderly people falls in smart homes is needed to increase the survival rate of the person or offer required support. Recently, the advent of artificial intelligence (AI), IoT, wearables, smartphones, etc. makes it feasible to design fall detection systems for smart homecare. In this view, this paper presents an IoT enabled elderly fall detection model using optimal deep convolutional neural network (IMEFD-ODCNN) for smart homecare. The goal of the IMEFD-ODCNN model is to enable smartphones and intelligent deep learning (DL) algorithms to detect the occurrence of falls in the smart home. Primarily, the input video captured by the IoT devices is pre-processed in different ways like resizing, augmentation, and min-max based normalization. Besides, SqueezeNet model is employed as a feature extraction technique to derive appropriate feature vectors for fall detection. In addition, the hyperparameter tuning of the SqueezeNet model takes place using the salp swarm optimization (SSO) algorithm. Finally, sparrow search optimization algorithm (SSOA) with variational autoencoder (VAE), called SSOA-VAE based classifier is employed for the classification of fall and non-fall events. Finally, in case of fall event detected, the smartphone sends an alert to the caretakers and hospital management. The performance validation of the IMEFD-ODCNN model takes place on UR fall detection dataset and multiple cameras fall dataset. The experimental outcomes highlighted the promising performance of the IMEFD-ODCNN model over the recent methods with the maximum accuracy of 99.76% and 99.57% on the multiple cameras fall and UR fall detection dataset.

INDEX TERMS Smart homecare, smartphone, fall detection, artificial intelligence, elderly people, deep learning, parameter tuning.

I. INTRODUCTION

In recent years, the Internet of Things (IoT) and mobile communication find useful in healthcare sector. With an enhanced healthcare system in several countries, average life span has developed considerably. Plus lower natural increases result in an elderly population that would need appropriate care and more interest. But, in several countries, offering appropriate care could be challenging because of several reasons. The impaired and elderly populations would shortly live in smart homes [1], [2]. These homes offer a pleasant and safe place for the elders. Independently, security is considering the main concern in the smart healthcare model [3]. However, daily emergency incidents will also continue to occur due to seniors' human nature. Falling is the most common problem encountered by elder peoples. For elder adults, a fall

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Hua Chen.

could be highly risky and might cause serious health issues. Additionally, lack of balance and fall might be symptoms of a life-threatening disease. Nevertheless of the cause for a fall, it can be critical if it happens, the injured people must obtain quick help. Frequently, the individual might not be able to rise up with no support and might require immediate medical consideration.

Unreported cases result in the fall of injury that may involve earlier treatments. Fear of falling increases the negative post fall effects and may decrease patient confidence [4]. Consequently, it limits the patient's activities, decreases social interaction, and finally causes depression [5], [6]. Respectively, it aids to decrease treatment costs and raise the opportunity of recovery. In [7], researchers have divided fall detection systems into 3 classes regarding cameras, wearable devices, and ambiance sensors. The system is depending upon wearable device seems to be common as they could identify a fall precisely nevertheless of the patient locations (viz., outdoor & indoor) and don't interrupt the person's privacy and day-to-day activities. Because of their asset limitations (for example storage capacity & limited power), it must have an innovative scheme that assists to decrease computation heavier loads on wearable sensor nodes, when preserving/enhancing the QoS.

A. NEED OF IoT ENABLED AI TECHNIQUES FOR FALL DETECTION

The independent life of elder persons could be altered significantly afterward a fall. Based on health state of the elders, nearly ten percent of the persons fall would endure severe injuries, or may even pass away straight afterward a fall when no intermediary aid is presented [8], [9]. For preventing the serious effects of this fall, consistent fall detection is required. The most popular method for detecting falls is wrist worn detection system which measures the acceleration force. These wrist devices are attaining more interest over the population and become gradually stronger based on computation efficiency that the utilization of AI is moderate. Generally, elder person appears to be attentive in utilizing these devices while they reveal concern on privacy and understand accurately when the device is processing at certain times [10]. Various fall detection methods have been presented in previous years. This method ranges from simple threshold based techniques, on handcrafted feature based ML technique, and lastly to DL based automated feature extraction NN.

IoT is the most appropriate candidate for this system since it contains broad innovative techniques like WSN, CC, and sensing to interconnect virtual objects using physical objects. As the gateways could execute difficult fall detection techniques like discrete wavelet transform/data mining. Additionally, smart gateway helps to enhance QoS by offering innovative services viz. local storage to store temporary data/push notification to inform anomaly in real world. It is predictable that IoT could widely assist in reducing power consumption of wearable devices with the allocation of tasks. But, IoT could not often assurance a higher level of energy efficacy in wearable devices. Another main problem is data transmission and data acquisition cause higher energy utilization in wearable sensor nodes should be considerately deliberated. If a wearable sensor node is energy ineffective, it might cause untrustworthiness and decreases QoS.

B. PAPER CONTRIBUTIONS

This paper presents an intelligent IoT enabled elderly fall detection model using optimal deep convolutional neural network (IMEFD-ODCNN) for smart homecare. At the initial stage, the input video captured by the IoT devices is pre-processed in different ways like resizing, augmentation, and min-max based normalization. Moreover, SqueezeNet model is used as a feature extractor and its hyperparameters are tuned by the use of SSO algorithm. Furthermore, sparrow search optimization algorithm (SSOA) with variational autoencoder (VAE), called SSOA-VAE based classifier is employed. The SSO algorithm is preferable owing to its high efficiency, robustness, accuracy, and convergence rate. The VAE is chosen because of the capability of learning smooth latent state representations of the input data. Lastly, in case of fall event detected, the smartphone sends an alert to the caretakers and hospital management. An extensive set of simulations is carried out on UR fall detection dataset and multiple cameras fall dataset. The key contribution of the paper is given as follows.

- Propose a novel IMEDF-ODCNN model for elderly fall detection in smart homecare
- Develop a hyperparameter tuned SqueezeNet based feature extractor with SSO algorithm to generate useful set of feature vectors
- Design an SSOA-VAE based classification model to detect the occurrence of fall and non-fall events
- Enables the smartphone to generate an alert to the caretakers and hospital authority on the occurrence of fall event
- Validate the fall detection performance of the IMEFD-ODCNN model on UR fall detection dataset and multiple cameras fall dataset.

C. PAPER ORGANIZATION

The rest of the paper is organized as follows. Section 2 briefs the existing fall detection approaches and section 3 describes the overall system architecture. Then, section 4 explains the different modules involved in the proposed IMEFD-ODCNN model. Next, section 5 assesses the experimental results and section 6 draws the concluding remarks.

II. LITERATURE REVIEW

Hussain *et al.* [11] presented a wearable sensor based continuous fall monitoring scheme that can detect falling and identify fall patterns and the activity related to fall incidents. The efficiency of the presented system is examined by a sequence of studies with 3 ML techniques as, RF, KNN, & SVM. Aziz *et al.* [12] examined the accuracy of fall detection

scheme deepening upon real time fall & non-fall datasets. The 5 younger & nineteen elder persons went on their everyday works when wearing tri axial accelerometer. Elderly persons suffered ten unexpected falls at the time of collecting data. Around four hundred hours of ADL have been noted. They utilized ML, SVM classification for identifying falls & non-fall activities. Shojaei-Hashemi *et al.* [13] proposed DL based method for detecting human fall, with the help of LSTM-NN. This module isn't limited to other certain conditions, and efficiency evaluation shows that it exceeds overall present techniques. Tsinganos and Skodras [14] developed a smartphone based fall detection scheme that could differentiate among ADL & falls. The usual fall detection scheme contains notification module and sensing component. Android devices, armed with communication services and sensors, are optimal candidates for the growth of this technique.

In Liu *et al.* [15], a sensing module combined energy efficient sensor was established that could sense and store the information of human activities from sleep mode, and interrupt driven technique is presented for transmitting the data to a server combined with Zigbee. Next, an FD-DNN operation on the server is designed carefully for detecting accurate falls. The FD-DNN integrated CNN alongside LSTM techniques was verified using offline & online datasets. In Kong *et al.* [16], a HOG-SVM based fall detection IoT scheme for elder adults was presented. For ensuring privacy and strong modifications of the light intensity, deep sensor is utilized rather than RGB camera for getting binary images of elder adults. Afterward attaining the denoised binary images, the features of person are extracted using the histogram of oriented gradient, and the image classification is executed to judge the fall condition using linear SVM.

Carletti *et al.* [17] proposed a new smartphone based fall detection scheme that considers falls as abnormalities regarding a module of usual events. This technique is related to other methods and it is demonstrated to be appropriate to operate on a smartphone located in the trouser pockets. This outcome is established from the attained accuracy and essential hardware assets. Mrozek *et al.* [1] proposed a scalable framework of a scheme that could observe 1000s of elder persons, identify falls, and inform the care takers. Scalability test discloses the need for enabling large scale scheme processes have been executed. Furthermore, they authenticated various ML modules for evaluating their appropriateness in the detection procedure. Amongst the tested modules, Boosted Decision Tree results in the optimal classification efficiency.

For improving the classification accuracy, the data from smartphones & smartwatches are integrated into [18]. There aren't various publicly available datasets integrating data from smartphones & smartwatches. Henceforth, the data would be independently gathering. The DT (J48) classification would be utilized for classifying the falls. Gia *et al.* [19] proposed the implementation of lightweight, tiny, energy efficient wearable, and flexible devices. Though several methods are available in the literature, it is needed to examine distinct variables (for example transmission protocol, communication bus interface, transmission rate, and sampling rate) impact on energy utilization of the wearable device. Additionally, they give complete analyses of energy utilization of the wearable in distinct configurations and operating situations. Also, it is give suggestions (software & hardware) for implementing an optimum wearable device to IoT based fall detection system based on higher QoS and energy efficiency.

III. SYSTEM ARCHITECTURE

The overall system architecture of the proposed model is depicted in Fig. 1. The proposed fall detection model uses a smartphone for processing. The IMEFD-ODCNN model allows smartphones and intelligent DL algorithms to detect the occurrence of falls in the smart home. The proposed IMEFD-ODCNN model involves distinct stages of operations like data acquisition, pre-processing, SqueezeNet based feature extraction, SSO based parameter tuning, and SSOA-VAE based classification. Primarily, the input videos are captured and are sent to the cloud server for additional processing where the proposed model gets executed.

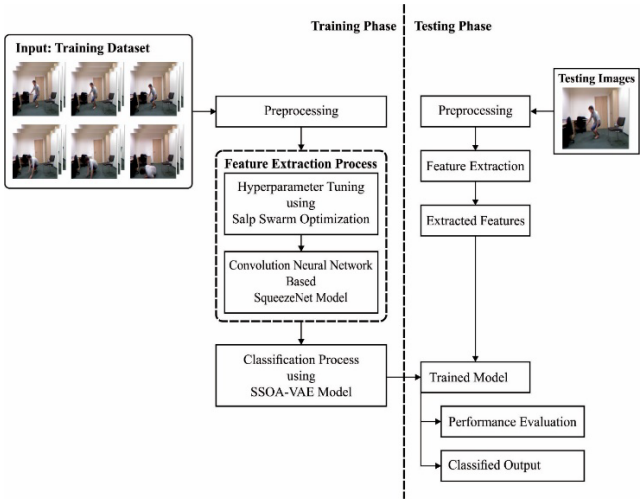

FIGURE 1. The working process of IMEFD-ODCNN model.

Then, the video frames are split and are pre-processed in three major levels such as resizing, augmentation, and normalization to enhance the quality of the video frames. Afterward, the features from the video frames are extracted to derive useful feature vectors using SqueezeNet model. Moreover, the hyperparameter tuning of the SqueezeNet model takes place using the SSO algorithm. Subsequently, the feature vectors are fed into the SSOA-VAE based classifier model to detect the occurrence of falls. Based on the classifier results, the subsequent actions will be performed. According to the value of classification outcome, the following actions are taken:

- When an event is detected as a fall and is denoted as class 1, an alarm is transmitted to the patient device from where the caretaker can be notified automatically if the fall was not excluded from the application by the monitored person.

- When an event is detected as non-fall event and is represented as class 0, no alarm will be transmitted and the event occurrence is discarded.

By the use of backend systems, the physicians/caretakers could observe the elderly people in real time from remote areas. Besides, the backend system aid doctors to treat diseases using the offered data and patient history.

IV. WORKING PROCESS OF IMEFD-ODCNN MODEL

The overall working process of the IMEFD-ODCNN model involves different subprocesses data acquisition, preprocessing, SqueezeNet based feature extraction, SSO based parameter tuning, and SSOA-VAE based classification. The detailed working of these processes is discussed in the succeeding subsections.

A. DATA PRE-PROCESSING

In the beginning stage, the frames were pre-processed for improving the characteristics of an image, removal the noise artefacts, and improve specific groups of features. At this point, the frames were processing from 3 important levels namely resizing, augmentation and normalization. In order to decrease the calculation cost, the resizing of frames occur from 150 × 150. At the same time, the frames are augmented where the frames are changed at all training epochs. For augmenting the frames, various models like zooming, horizontal flipping, rotation, width, and height shifting. At last, normalization technique was implemented to enhanced generalization of the model.

B. SqueezeNet BASED FEATURE EXTRACTION

The CNN generally contains full connection layer, convolutional layer, and pooling layer. Initially, the feature is extracted with more than one pooling & convolution layer. Later, entire feature mappings from the latter convolution layer are converted to 1D vectors for full connection. Lastly, the output layer categorizes the input images. The network alters the weight variables using BP and minimizes the square variance among the classification outcomes and predictable output. The neurons in every layer are ordered in 3D: depth, width, and height, where height & width is the size of neuron, and depth denotes channel amount of the input image/the amount of input feature mappings. The convolutional layer has many convolution filters, extract distinct features from the image using convolution process. The convolution filter of the present layer convoluted the input feature mappings for extracting local features and attain the output feature mappings. Later, the nonlinear feature mappings are attained with activation function. The pooling layer, so called subsampling layer, is behindhand the convolutional layer. It executes down sampling process, with a certain value as output in specific regions. With the removal of insignificant instance points from the feature map, the size of input feature map of the following layer is decreased, and the computation complexity is also reduced. Simultaneously, the flexibility of the network to the modifications of image rotation & translation was also raised [20]. The general pooling operation contains average and maximal pooling. The framework is denuding upon pooling & convolutional layers could enhance the strength of the network module. The CNN could expand by multilayer convolutions. By amount of increasing layers, the features attained via learning becomes global. Eventually, the global feature map learned is converted to a vector for connecting full connection layer. All variables in the network module are in the full connection layer.

Since the number of variables for VGGNet & AlexNet is increasing, the SqueezeNet network module was presented that has minimal variables when maintaining accuracy. The fire model is the fundamental model in SqueezeNet, and its structure is displayed in Fig. 2. This model is separated to Expand & Squeeze frameworks. The 1 × 1 convolutional layer has gained more interest in the deliberation of network structure. The works explain from the perception of cross channel pooling where MLP is equal to the cascade cross channel parametric pooling layer behindhand the conventional kernel, therefore attaining a linear integration of multiple feature maps and data incorporation over the channels. If the number of output & input channels are larger, the convolution kernel variable becomes larger. They include 1 × 1 convolution for all inception modules, decreasing the amount of input channels, and the convolution kernel variables and complexity operation is reduced. Finally, a 1 × 1 convolution is included for improving the number of channels and improve feature extraction. If the sampling reduction process is delayed, a large activation graph is given to the convolutional layer, whereas the large activation graph maintains additional data that could give high classification accuracy [27]–[29].

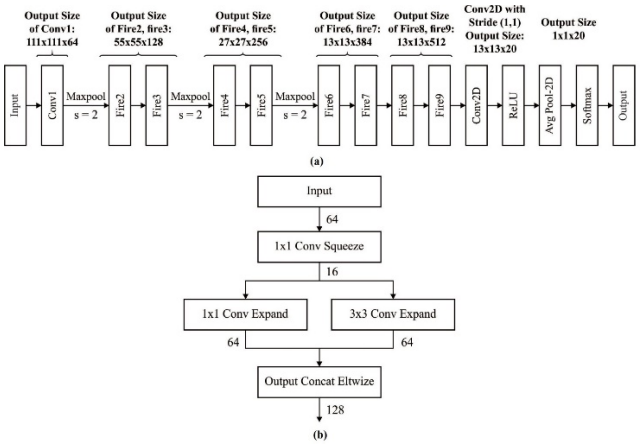

FIGURE 2. a) SqueezeNet structure (b) Fire module with layers.

C. HYPERPARAMETER OPTIMIZATION USING SSO ALGORITHM

For tuning the hyperparameters of the SqueezeNet model, the SSO algorithm is applied to optimally adjust the

hyperparameters involved in it. The salps included the set of Salpidae that comprises a visible barrel shaped body. The tissues are identical to jellyfishes. As well, the motion is identical to jelly fish, when the water is inspired by a body as propulsion and goes in the forward direction. A mathematical depiction of swarming behaviors & population of salps is determined. In addition, no mathematical method of salp swarm is utilized for resolving optimization problems were swarms of fishes, bees, and ants are widely applied and labeled to solve the enhanced problem. For modeling the salp chain mathematically, the population is categorized by two classes such as Follower and Leader. Firstly, leader is considered to be salp at the front phase of a chain, whereas the residual salp is so called follower. According to the names, the salps represents leader guide the swarm where the follower follows each other. Compared with other swarm based modules, the place of salps is determined as n-dimension search space where n represents variable amount of the employed problem. Henceforth, the position of salps is kept in a two dimensional matrix and so called x. Furthermore, it consider a food source so called F whereas search space is the swarm target. For upgrading the leader position, it is represented by.

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0 \\ F_j - c_1((ub_j - 1b_j)c_2 + 1b_j) & c_3 < 0 \end{cases} \quad (1)$$

where $as\, x_j^1$ denotes location of primary salp in jth dimension, F_j represents position of food source in jth dimension, ub_j indicates maximal bound of jth dimension, lb_j denotes minimal bound of jth dimension, c_1, c_2, and c_3 denotes random values [21].

Eq. (1) represents leader is preferred to update the location regarding food sources. The coefficient c_1 is more important feature in SSA since it deals with the exploitation and exploration as determined by:

$$c_1 = 2e^{-(\frac{41}{L})^2} \quad (2)$$

whereas l denotes current iteration where L indicates high amount of iteration. The attribute c_2 and c_3 are determined as random values which are generated uniformly with zero and one. The forthcoming location in jth dimension is negative/positive infinity and step size. To upgrade the position of follower, the given function is employed:

$$\mathrm{x_j^i} = \frac{1}{2}\mathrm{at}^2 + v_0 t \quad (3)$$

whereas $i \geq 2, x_j^i$ denotes position of ith follower salp in jth dimension, t indicates time, v_0 represents primary speed, and $a = \frac{v_{final}}{v_0}$ where $v = \frac{x - x_0}{r}$. Because the time in optimization denotes iteration, the difference between iterations is one, and assume that $v_0 = 0$, whereas the functions are employed by:

$$x_j^i = \frac{1}{2}\left(x_j^i + x_j^{i-1}\right) \quad (4)$$

whereas $i \geq 2$ and x_j^i denotes location of ith follower salp in jth dimension. By Eqs. (1) and (4), salp chain can be speeded.

D. FALL DETECTION USING SSOA-VAE MODEL

During the classification stage, the SSOA-VAE model gets executed to determine the class labels of the input video frames, i.e. non-fall or fall event. A VAE is a variation of AE rooted in Bayesian inference. It can module the fundamental distribution of observation p (z) and generates novel data by presenting a group of latent arbitrary parameters z. They could denote the procedure as $p(\mathrm{x}) = \int p(\mathrm{x}|\mathrm{z})\, p(\mathrm{z})\, d\mathrm{z}$. But, the marginalization is computationally intractable as the search space of z is constant and combinatorically larger. Instead, they could denote marginal log probability of a separate points $\log p(\mathrm{x}) = D_{KL}(q_\varphi(z|x)||p_\theta(z)) + \mathcal{L}_{\mathrm{vae}}(\varphi, \theta; x)$ with representation from [22], whereas D_{KL} denotes Kullback Leibler divergence from previous $p_\theta(\mathrm{z})$ to the variation calculation $q_\varphi(\mathrm{z}|\mathrm{x})$ of $p(\mathrm{z}|\mathrm{x})$ and $\mathcal{L}_{\mathrm{vae}}$ indicates variation lower bound of the data x with Jensen's inequality. Noted that φ and θ denote variables of the encoder & decoder, correspondingly. Fig. 3 demonstrates the structure of VAE.

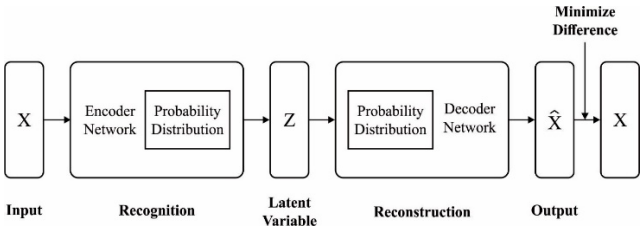

FIGURE 3. Architecture of VAE.

A VAE enhances the variables, φ, and θ, with the maximization of lower bound of the log probability, $\mathcal{L}_{\mathrm{vae}}$,

$$\mathcal{L}_{\mathrm{vae}} = -D_{KL}(q_\varphi(z|x)||p_\theta(z)) + E_{q_\varphi}(z|x)[logp_\theta(x|z)]. \quad (5)$$

The initial word standardizes the latent parameter z with the minimization of KL divergence among estimated previous and posterior of the latent parameter. Next the recreation of x with the maximization of log probability $\log p_\theta(\mathrm{x}|\mathrm{z})$ by sampling from $q_\varphi(\mathrm{z}|\mathrm{x})$. The selection of distribution kinds is significant as VAE modules the estimated posterior distribution $q_\varphi(z|x)$ from previous $p_\theta(z)$ and probability $p_\theta(\mathrm{x}|\mathrm{z})$. A usual selection for the posterior is Gaussian distribution, $\mathcal{N}(\mu_z, \Sigma_z)$, whereas a typical standard distribution $\mathcal{N}(0, 1)$ is utilized prior. For the probability, a Bernoulli distribution/multivariate Gaussian distribution is frequently utilized to continuous/binary data, correspondingly.

In order to determine the parameters involved in the VAE (φ and θ), the SSOA is applied in such a way that the fall detection performance gets improved. The sparrows are usually kindly birds and contain several types. They are distributed worldwide and attentive to survive in areas around humans. Similarly, they are omnivorous species and mostly eat seeds and grains. It is usually so called resident in nature. In comparison with other little birds, it is stronger in memory power and creativeness. It has two different types of captive house sparrows such as scrounger and producer. The producers strongly search for the food source, while the scrounger

acquires foodstuff from the producer. Furthermore, the proof exhibits that the bird usually exploits behavioural approach adaptably, and shift among scrounger & producer. From the study, it is exposed that the sparrow finds their food by the approach of producer and the scrounger based on circumstances. It is noteworthy that the birds are placed on the edge of the population, are possible to be attacked through predators, and continuously try to attain an optimal location. The sparrows are located on the central may travel to their neighbours for reducing the threat.

Initially, the virtual sparrow is employed to identify an optimum source of food. The residence of sparrow is given by:

$$X = \begin{bmatrix} \chi_{1,1} & \chi_{1,2} & \cdots & \cdots & \chi_{1'd} \\ \chi_{2,1} & \chi_{2,2} & \cdots & \cdots & \chi_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & \cdots & x_{n,d} \end{bmatrix} \tag{6}$$

whereas n denotes amount of sparrow and d indicates direction of variable that should be enhanced. Therefore, the fitness score of all sparrows are defined by [23]:

$$F_X = \begin{bmatrix} f([x_{1,1} & \chi_{1,2} & \cdots\cdots & \chi_{1,d}]) \\ f([x_{2,1} & \chi_{2,2} & \cdots\cdots & \chi_{2,d}]) \\ \vdots & \vdots & \vdots\ \vdots & \vdots \\ f([x_{n,1} & x_{n,2} & \cdots\cdots & x_{n,d}]) \end{bmatrix} \tag{7}$$

whereas n denotes amount of sparrow, and calculate of entire rows in F_X represents fitness score of an individual. In SSA, the producer contains maximal fitness measure that accomplishes an optimum food in search function. Likewise, producer is responsible for exploring food and support the action of entire populations. Hence, the producer can detect food in wide range than scroungers. According to rules (6) and (7), the location of producer is expanded by:

$$X_{j,j}^{t+1} = \begin{cases} X_{i,j}^{t} \cdot \exp(\frac{-i}{\alpha \cdot iter_{\max}}) & if\ R_2 < ST \\ X_{i,j}^{t} + O \cdot L & if\ R_2 \geq ST \end{cases} \tag{8}$$

whereas t indicates current iteration, $j = 1, 2, \ldots, d.X_{j,j}^{t}$ denotes rate of jt dimension of it sparrow at iteration $t.iter_{\max}$ represents constant with maximal iteration. $\alpha \in (0, 1]$ denotes random value. $R_2(R_2 \in [0, 1])$ and $ST(ST \in [0.5, 1.0])$ represents an alarm value and security threshold respectively. Q denotes arbitrary value that employs simple distribution and L indicates matrix $1 \times d$ for entire components with one.

While $R_2 < ST$, denotes no predator exists, and producers get to wide search mode. Once $R_2 \geq ST$, then some sparrows have created the predator, and it is important to protect them by flying to safe region.

For scrounger, it employs rules (9) and (10). Few scroungers follow the producer obviously. When the producer finds an optimal food, then it leaves the location for competing to food. When the competitions are effective, then they could attain the food of producer, or rules (10) is executed. The position upgrades formal for scrounger is given by:

$$X_{i,j}^{t+1} = \begin{cases} 0.\exp(\frac{X_{worst}^{t} - X_{i}^{t_i}}{j^2}) & if\ i > n/2 \\ X_{P}^{t+1} + |X_{j,j}^{t} - X_{P}^{t+1}| \cdot A^{+} \cdot L & otherwise \end{cases} \tag{9}$$

whereas X_P denotes optimum location employed with a producer, X_{worst} indicates current global worst location, A showcase a matrix of $1 \times d$ for a component with one, and $A^{+} = A^{T}(AA^{T})^{-1}$. If $i > n/2$, it recommends that it scrounger with unsuccessful fitness is highly starving.

Consequently, the sparrows are away from danger will contain further lifespan. The main position of sparrow is created arbitrarily in the population. Depending upon, the arithmetical technique is given by:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \cdot |X_{i,j}^{t} - X_{best}^{t}| & if\ f_i > f_g \\ X_{i,j}^{t} + K \cdot \left(\frac{|X_{i,j}^{t} - X_{worst}^{t}|}{(f_j - f_w) + \varepsilon} \right) & if\ f_i = f_g \end{cases} \tag{10}$$

whereas X_{best} denotes current global optimum position. β, indicate step size control variable that is a standard distribution of random values with mean value of zero and a variance of one. $K \in [-1, 1]$ denotes arbitrary measure. In this module, f_i denotes fitness value of current sparrow. f_g and f_w indicates current global optimum and worst fitness measures. ε indicate minimal constant and remove zero division error.

In event of easiness, when $f_i > f_g$ determines sparrow is at edge of a group. X_{est} indicates location of a centre of population that is secure. Now, $f_j = f_g$ indicates sparrow, is in the center of a population which is attentive from danger and travels nearer to the border. K denotes direction whereas sparrow moves and step size control coefficient.

V. PERFORMANCE VALIDATION

The proposed model is validated using Multiple cameras fall dataset [24] and UR Fall Detection (URFD) dataset [25]. The first dataset comprises 192 videos where 96 videos come under fall events and 96 videos come under non-fall events. The second dataset has frontal sequence of 314 frames, where 74 frames come into fall event and 240 frames come under non-fall event. The parameter setting of the proposed model is given as follows: mini batch size: 200, dropout: 0.5, number of hidden layers:3, and number of hidden units: 1024.

Fig. 4 shows the sample test images from the dataset. Besides, the depth level of the images from the sample test images is illustrated in Fig. 5.

The classification results analysis of the IMEFD-ODCNN model on multiple cameras falls dataset is given in Table 1 under varying training size (TS). On the TS of 40%:60%, the IMEFD-ODCNN model has gained a specificity of 99.10%, precision of 99.67%, recall of 99.82%, accuracy of 99.54%, and F-score of 99.46%. Moreover, on the TS of 60%:40%, the IMEFD-ODCNN technique has achieved a specificity of 99.22%, precision of 99.51%, recall of 99.81%, accuracy of 99.80%, and F-score of 99.53%. Furthermore, on the TS of 80%:20%, the IMEFD-ODCNN methodology has obtained a

TABLE 1. Comparative analysis of proposed IMEFD-ODCNN method on multiple cameras fall dataset in terms of different measures.

Training Size	Specificity	Precision	Recall	Accuracy	F-score
40%:60%	99.10	99.67	99.82	99.54	99.46
50%:50%	99.12	99.45	99.85	99.62	99.49
60%:40%	99.22	99.51	99.81	99.80	99.53
70%:30%	99.45	99.76	99.95	99.89	99.87
80%:20%	99.56	100	99.83	99.94	99.92
Average	99.29	99.68	99.85	99.76	99.65

FIGURE 4. Sample sequences.

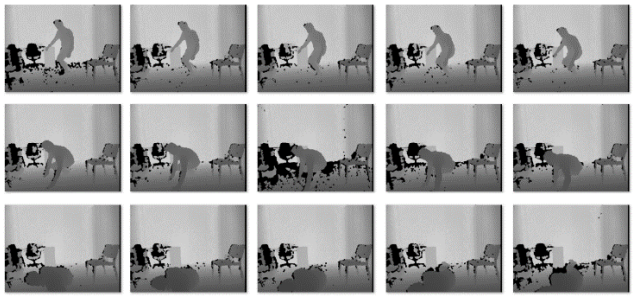

FIGURE 5. Depth data of the sequences.

specificity of 99.56%, precision of 100%, recall of 99.83%, accuracy of 99.94%, and F-score of 99.92%.

Fig. 6 illustrates the ROC analysis of the IMEFD-ODCNN model on the multiple cameras fall dataset under different TS. From the figure, it is evident that the IMEFD-ODCNN model has accomplished effective fall detection performance with the maximum ROC values under all TS. The classification outcomes analysis of the IMEFD-ODCNN approach on UR Fall Detection dataset is provided in Table 2 under varying training size (TS). On the TS of 40%:60%, the IMEFD-ODCNN method has attained a specificity of 99.09%, precision of 99.93%, recall of 99.47%, accuracy of 99.49%, and F-score of 99.32%.

Furthermore, on the TS of 60%:40%, the IMEFD-ODCNN method has reached a specificity of 99.76%, precision of 99.81%, recall of 99.59%, accuracy of 99.54%, and F-score of 99.30%. Also, on the TS of 80%:20%, the IMEFD-ODCNN technique has achieved a specificity of 99.89%, precision of 99.56%, recall of 99.90%, accuracy of 99.87%, and F-score of 99.59%.Fig. 7 showcases the ROC analysis of the IMEFD-ODCNN method on the UR Fall Detection Dataset in distinct TS. From the figure, it can be stated that the IMEFD-ODCNN technique has accomplished effective fall detection performance with the maximal ROC values in all TS.

A brief comparison study of the IMEFD-ODCNN model with other existing methods on multiple cameras fall dataset take place in Table 3. From the figure, it is demonstrated that the 1D Conv NN and 2D Conv NN models have demonstrated poor results with the accuracy of 94.3% and 95.5% respectively. In line with that, the ResNet-50 and ResNet-101 models have accomplished moderately closer outcomes with the accuracy of 96.1% and 96.5% respectively. Next to that, the Depthwise, VGG-16, and VGG-19 models have accomplished moderately closer outcomes with the accuracy of 97.8%, 98%, and 98% respectably. Finally, the proposed IMEFD-ODCNN model has showcased superior results with an accuracy of 99.76%.

Another comparative study of the IMEFD-ODCNN methodology with other state-of-art techniques on UR Fall Detection Dataset takes place in Table 4. From the figure, it can be outperformed that the 1D Conv NN and 2D Conv NN techniques have showcased worse outcomes with the accuracy of 92.7% and 95% correspondingly. Along with that, the ResNet-50 and ResNet-101 manners have accomplished moderately closer result with the accuracy of 95.4% and 96.2% correspondingly. Likewise, the VGG-16, Depthwise, and VGG-19 algorithms have accomplished moderately closer result with the accuracy of 97.6%, 98%, and 98% correspondingly. Eventually, the projected IMEFD-ODCNN technique has outperformed maximum outcomes with an accuracy of 99.57%.

An extensive training time and testing time analysis of the proposed IMEFD-ODCNN model with other existing techniques on Multiple Cameras Fall dataset is given in Table 5 and Fig. 8. From the results, it is evident that the VGG-16 model has required maximum training and testing time of 3627.48s and 1758.33s respectively. At the same time, the VGG-19 model has needed a slightly reduced training and testing time of 3189.03s and 1482.46s respectively. Besides, the ResNet-101 model has resulted in a moderate training and testing time of 1274.13s and 932.4s respectively. Meanwhile, the 2D Conv NN and 1D Conv NN models have showcased moderate testing and training time. Eventually, the ResNet-50 model has needed a competitively decreased training and

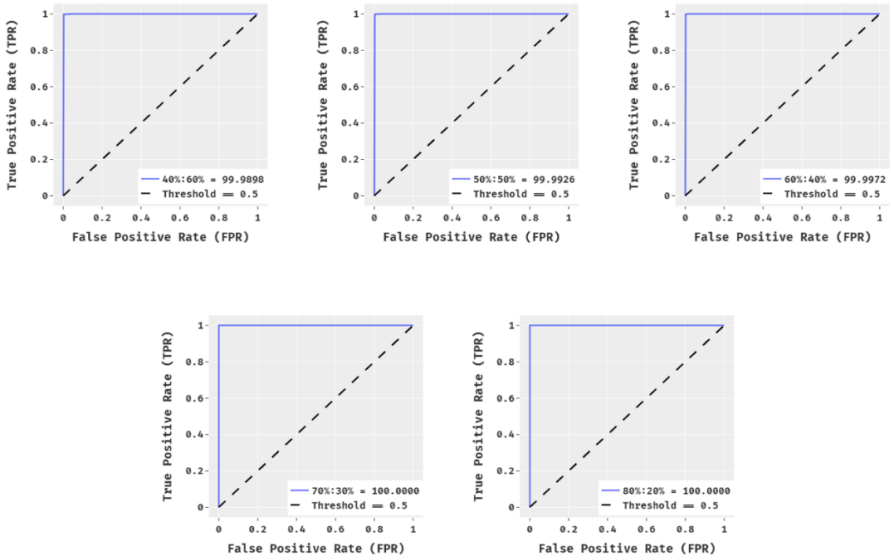

FIGURE 6. ROC analysis of proposed IMEFD-ODCNN method on multiple cameras fall dataset.

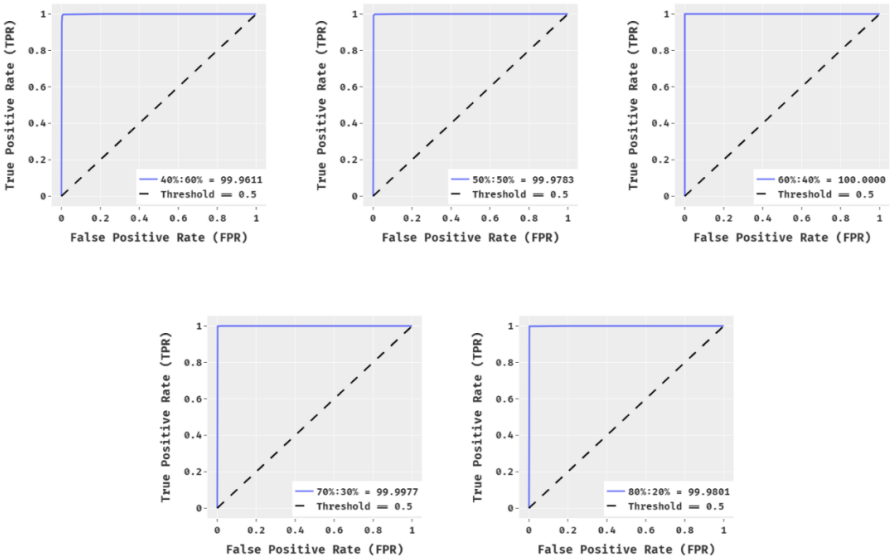

FIGURE 7. ROC analysis of proposed IMEFD-ODCNN method on UR fall detection dataset.

TABLE 2. Comparative analysis of proposed IMEFD-ODCNN method on UR fall detection dataset in terms of different measures.

Training Size	Specificity	Precision	Recall	Accuracy	F-score
40%:60%	99.09	99.93	99.47	99.49	99.32
50%:50%	99.23	99.38	99.21	99.43	99.19
60%:40%	99.76	99.81	99.59	99.54	99.30
70%:30%	99.61	99.90	99.58	99.52	99.52
80%:20%	99.89	99.56	99.90	99.87	99.59
Average	99.52	99.72	99.55	99.57	99.38

testing time of 1163.45s and 946.22s respectively. However, the proposed IMEFD-ODCNN model has showcased effective outcomes with the training and testing time of 1136.94s and 735.41s respectively.

An extensive training time and testing time analysis of the projected IMEFD-ODCNN technique with other existing methods on UR Fall Detection Dataset is provided in Table 6 and Fig. 9 [26]. From the results, it can be revealed

TABLE 3. Comparative analysis of existing with proposed IMEFD-ODCNN method on multiple cameras fall dataset in terms of accuracy.

Methods	Accuracy (%)
VGG-16 Model	98.00
VGG-19 Model	98.00
Depthwise Model	97.80
1D Conv NN	94.30
2D Conv NN	95.50
ResNet-50 Model	96.10
ResNet-101 Model	96.50
IMEFD-ODCNN	99.76

TABLE 4. Comparative analysis of existing with proposed IMEFD-ODCNN method on UR fall detection dataset in terms of accuracy.

Methods	Accuracy (%)
VGG-16 Model	97.60
VGG-19 Model	98.00
Depthwise Model	98.00
1D Conv NN	92.70
2D Conv NN	95.00
ResNet-50 Model	95.40
ResNet-101 Model	96.20
IMEFD-ODCNN	99.57

TABLE 5. Comparative analysis of existing with proposed IMEFD-ODCNN method on multiple cameras fall dataset in terms of training time and testing time.

Methods	Training Time (sec)	Testing Time (sec)
VGG-16	3627.48	1758.33
VGG-19	3189.03	1482.46
Depthwise Model	2747.18	948.34
1D Conv NN	2341.40	924.58
2D Conv NN	1903.44	845.64
ResNet-50 Model	1163.45	946.22
ResNet-101 Model	1274.13	932.40
IMEFD-ODCNN	1136.94	735.41

that the VGG-16 technique has required higher training and testing time of 2352.6s and 1108.8s correspondingly. Simultaneously, the VGG-19 approach has needed a somewhat diminished training and testing time of 2778.6s and 1372.2s respectively.

Also, the ResNet-101 method has resulted in a moderate training and testing time of 1545.6s and 925.8s correspondingly. In the meantime, the ResNet-50 and 2D Conv NN models have showcased moderate testing and training time. Finally, the Depthwise model has needed a competitively reduced training and testing time of 1093.2s and 725.4s respectively. However, the presented IMEFD-ODCNN methodology has outperformed effective results with the training and testing time of 1014s and 677.4s correspondingly. The experimental outcomes highlighted the

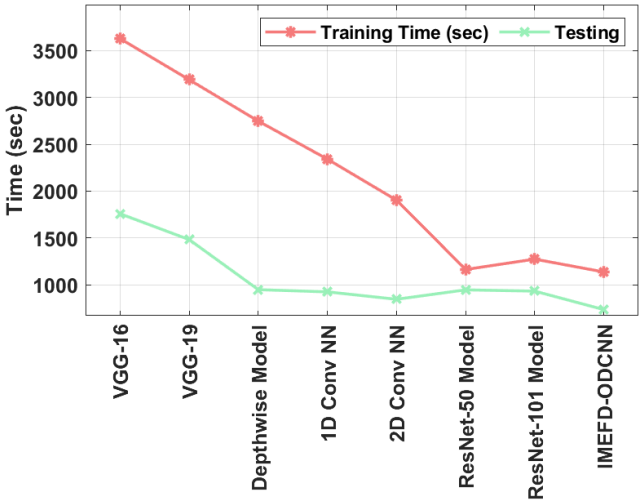

FIGURE 8. Comparative analysis of IMEFD-ODCNN model on multiple cameras fall dataset.

TABLE 6. Comparative analysis of existing with proposed IMEFD-ODCNN method on UR fall detection dataset in terms of training time and testing time.

Methods	Training Time (sec)	Testing Time (sec)
VGG-16	2352.60	1108.80
VGG-19	2778.60	1372.20
Depthwise Model	1093.20	725.40
1D Conv NN	1173.60	828.00
2D Conv NN	1228.80	780.00
ResNet-50 Model	1420.80	879.00
ResNet-101 Model	1545.60	925.80
IMEFD-ODCNN	1014.00	677.40

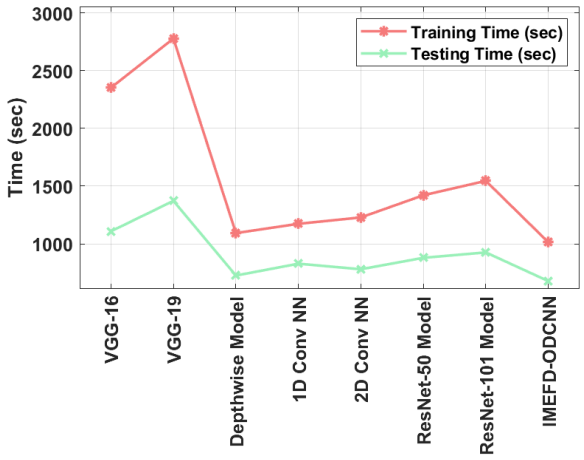

FIGURE 9. Comparative analysis of IMEFD-ODCNN model on UR fall detection dataset.

promising performance of the IMEFD-ODCNN model over the recent methods with the maximum accuracy of 99.76% and 99.57% on the multiple cameras fall and UR fall detection dataset. The proposed model outperforms the existing

methods due to the inclusion of SqueezeNet model and hyperparameter optimization process using SSOA.

VI. CONCLUSION

This paper has designed a new IMEFD-ODCNN model to detect fall events in smart homecare of elderly people. The IMEFD-ODCNN model allows IoT devices and intelligent DL algorithms to detect the occurrence of falls in the smart home. The proposed IMEFD-ODCNN model involves different stages of operations such as data acquisition, preprocessing, SqueezeNet based feature extraction, SSO based parameter tuning, and SSOA-VAE based classification. Once the fall is identified, an immediate alert is sent to the caretakers and hospital management. The utilization of SSO algorithm to select the hyperparameters of the SqueezeNet model and SSOA algorithm for parameter adjustments of the VAE model helps to considerably improve the overall fall detection performance. An extensive set of simulations is carried out on UR fall detection dataset and multiple cameras fall dataset. The experimental results highlighted the promising performance of the IMEFD-ODCNN model over the recent state of art methods. In future, the fall detection performance of the IMEFD-ODCNN model can be improved by the use of advanced DL models for classification process. Besides, scalable and robust versions of the IMEFD-ODCNN model can be developed to assist real time fall detection events from low-quality videos.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated during the current study.

REFERENCES

[1] D. Mrozek, A. Koczur, and B. Małysiak-Mrozek, "Fall detection in older adults with mobile IoT devices and machine learning in the cloud and on the edge," *Inf. Sci.*, vol. 537, pp. 132–147, Oct. 2020.
[2] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Phys. A, Stat. Mech. Appl.*, vol. 541, Mar. 2020, Art. no. 123344.
[3] B. Aguiar, T. Rocha, J. Silva, and I. Sousa, "Accelerometer-based fall detection for smartphones," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2014, pp. 1–6.
[4] C. H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. 103, no. 1, pp. 265–267, 2020.
[5] A. C. Scheffer, M. J. Schuurmans, N. van Dijk, T. van der Hooft, and S. E. D. Rooij, "Fear of falling: Measurement strategy, prevalence, risk factors and consequences among older persons," *Age Ageing*, vol. 37, no. 1, pp. 19–24, Jan. 2008.
[6] M. Pan, Y. Liu, J. Cao, Y. Li, C. Li, and C.-H. Chen, "Visual recognition based on deep learning for navigation mark classification," *IEEE Access*, vol. 8, pp. 32767–32775, 2020.
[7] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Trans. Inf. Syst.*, vol. E102.D, no. 7, pp. 1374–1383, Jul. 2019.
[8] J. A. Stevens, P. S. Corso, E. A. Finkelstein, and T. R. Miller, "The costs of fatal and non-fatal falls among older adults," *Injury Prevention*, vol. 12, no. 5, pp. 290–295, 2006.
[9] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4231–4232, Oct. 2018.
[10] S. Chaudhuri, H. Thompson, and G. Demiris, "Fall detection devices and their use with older adults: A systematic review," *J. Geriatric Phys. Therapy*, vol. 37, no. 4, pp. 178–196, Oct. 2014.
[11] F. Hussain, F. Hussain, M. Ehatisham-ul-Haq, and M. A. Azam, "Activity-aware fall detection and recognition based on wearable sensors," *IEEE Sensors J.*, vol. 19, no. 12, pp. 4528–4536, Jun. 2019.
[12] O. Aziz, J. Klenk, L. Schwickert, L. Chiari, C. Becker, E. J. Park, G. Mori, and S. N. Robinovitch, "Validation of accuracy of SVM-based fall detection system using real-world fall and non-fall datasets," *PLoS ONE*, vol. 12, no. 7, Jul. 2017, Art. no. e0180318.
[13] A. Shojaei-Hashemi, P. Nasiopoulos, J. J. Little, and M. T. Pourazad, "Video-based human fall detection in smart homes using deep learning," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
[14] P. Tsinganos and A. Skodras, "A smartphone-based fall detection system for the elderly," in *Proc. 10th Int. Symp. Image Signal Process. Anal.*, Sep. 2017, pp. 53–58.
[15] L. Liu, Y. Hou, J. He, J. Lungu, and R. Dong, "An energy-efficient fall detection method based on FD-DNN for elderly people," *Sensors*, vol. 20, no. 15, p. 4192, Jul. 2020.
[16] X. Kong, Z. Meng, N. Nojiri, Y. Iwahori, L. Meng, and H. Tomiyama, "A HOG-SVM based fall detection IoT system for elderly persons using deep sensor," *Procedia Comput. Sci.*, vol. 147, pp. 276–282, Jan. 2019.
[17] V. Carletti, A. Greco, A. Saggese, and M. Vento, "A smartphone-based system for detecting falls using anomaly detection," in *Proc. Int. Conf. Image Anal. Process.* Cham, Switzerland: Springer, Sep. 2017, pp. 490–499.
[18] S. Suresh, M. Jain, and R. Ramadoss, "Fall classification based on sensor data from smartphone and smartwatch," *AIP Conf. Proc.*, vol. 2112, no. 1, Jun. 2019, Art. no. 020075.
[19] T. Nguyen Gia, V. K. Sarker, I. Tcarenko, A. M. Rahmani, T. Westerlund, P. Liljeberg, and H. Tenhunen, "Energy efficient wearable sensor node for IoT-based fall detection systems," *Microprocessors Microsyst.*, vol. 56, pp. 34–46, Feb. 2018.
[20] G. Xu, "Deep convolutional neural network to detect J-UNIWARD," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2017, pp. 67–73.
[21] B. S. Ahmed, M. A. Sahib, and M. Y. Potrus, "Generating combinatorial test cases using simplified swarm optimization (SSO) algorithm for automated GUI functional testing," *Eng. Sci. Technol., Int. J.*, vol. 17, no. 4, pp. 218–226, Dec. 2014.
[22] D. Park, Y. Hoshi, and C. C. Kemp, "A multimodal anomaly detector for robot-assisted feeding using an LSTM-based variational autoencoder," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1544–1551, Jul. 2018.
[23] W. Tuerxun, X. Chang, G. Hongyu, J. Zhijie, and Z. Huajian, "Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm," *IEEE Access*, vol. 9, pp. 69307–69315, 2021.
[24] E. Auvinet, C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Multiple cameras fall dataset," DIRO-Université de Montréal, Montreal, QC, Canada, Tech. Rep. 1350, Jul. 2010.
[25] [Online]. Available: http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html
[26] A. Sultana, K. Deb, P. K. Dhar, and T. Koshiba, "Classification of indoor human fall events using deep learning," *Entropy*, vol. 23, no. 3, p. 328, Mar. 2021.
[27] B. A. Y. Alqaralleh, S. N. Mohanty, D. Gupta, A. Khanna, K. Shankar, and T. Vaiyapuri, "Reliable multi-object tracking model using deep learning and energy efficient wireless multimedia sensor networks," *IEEE Access*, vol. 8, pp. 213426–213436, 2020, doi: 10.1109/ACCESS.2020.3039695.
[28] A. Rajagopal, A. Ramachandran, K. Shankar, M. Khari, S. Jha, Y. Lee, and G. P. Joshi, "Fine-tuned residual network-based features with latent variable support vector machine-based optimal scene classification model for unmanned aerial vehicles," *IEEE Access*, vol. 8, pp. 118396–118404, 2020, doi: 10.1109/ACCESS.2020.3004233.
[29] V. Porkodi, A. R. Singh, A. R. W. Sait, K. Shankar, E. Yang, C. Seo, and G. P. Joshi, "Resource provisioning for cyber–physical–social system in cloud-fog-edge computing using optimal flower pollination algorithm," *IEEE Access*, vol. 8, pp. 105311–105319, 2020, doi: 10.1109/ACCESS.2020.2999734.