Received November 21, 2018, accepted December 21, 2018, date of publication January 17, 2019, date of current version March 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2893202

Multi Modal Aptitude Detection System for Smart Office

MOHAMMAD TEHSEEN[1], HUMA JAVED[1], AMJAD MEHMOOD[2], (Senior Fellow, IEEE), MUHAMMAD AMIN[3], ISHTIAQ HUSSAIN[1], AND BILAL JAN[4]

[1]Department of Computer Science, University of Peshawar, Peshawar, Pakistan
[2]Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan
[3]Department of Computer Science, Foundation for Advancement of Science and Technology, National University of Computer and Emerging Sciences, Peshawar, Pakistan
[4]Department of Computer Science, FATA University, Kohat, Pakistan

Corresponding author: Amjad Mehmood (dramjad.mehmood@ieee.org)

ABSTRACT Smart offices are places that aim to provide a favorable, interactive, and healthy environment for the employees. As a result, their working efficiency is enhanced. In order to achieve these numerous systems such as lie detector, emotion detection, and recognition, automation of office equipment has been proposed and implemented. To contribute toward one of the goals of smart offices, a novel ''multi-modal aptitude detection system'' is being proposed. The proposed system helps in analyzing the working aptitude of employees in order to help them in providing the more conducive and healthy working environment. In this paper, a detailed layered framework, implementation details, and achieved outcomes related to the proposed novel system are presented. The results achieved show that this novel system after successfully becoming a part of smart office can bring significant enhancements in the efficiency of employees. The proposed system helps in learning and identifying job preferences of employees based on their working aptitude. The system is validated and showed an accuracy of 96% in determining an employee's aptitude. The results are further verified by computing the F1 scores, and more details regarding testing and validation are provided.

INDEX TERMS Wireless sensor networks, wireless sensor based applications, classification, machine learning, signal processing, neural networks.

I. INTRODUCTION

Smart Offices are established to integrate physical devices, human beings and computing technologies with the intention of providing a healthy, conducive, interactive and intelligent environment for employees. These places are constantly evolving to provide more interconnected and motivated environment for planning and administering daily activities of employees at work [1]. Smart office's main objective is to enhance efficiency of employees, to achieve it different types of smart systems has been proposed and implemented [2]. In smart systems, sensors aligned with actuators (mechanisms that controls other devices) helps in facilitating employees to perform assigned tasks by automatically monitoring and controlling devices present in smart environment [3].

Wireless sensor based systems and networks are the most common approach followed to realize smart systems. These systems are emerging very rapidly as control and monitoring subsystems, for variety of applications such as Structural Monitoring [1], Industrial Process Control [3], Smart Offices [2], [3], Military or Border Surveillance [4], [5], Environmental Conditions Monitoring [6], Intelligent Agriculture [7], Health Care [8], [9], Pothole Detection [10], Home Intelligence [11], Smart Cities [12], Earth Quake Monitoring [13], Infant Monitoring [14], Lie Detection and Emotion Recognition [15], [16], Automatic Traffic Control [17], Internet of Things [18], etc.

Recently, emotion detection and recognition systems [18] are emerging as a valuable asset for smart spaces. In smart environments, these detection systems are used to detect different emotions, for example level of fatigue, frustration, happiness, sadness, in order to help people becoming more productive to smart environments. One such application is deployed in jails [17], which tries to determine the violent or aggressive behavior of prisoners to safeguard jail officials and other prisoners. In smart offices these systems are deployed to identify, whether an employee is under stress or feeling happiness or sad or feeling comfortable while performing his routine tasks.

However, these emotions are very difficult to measure accurately or even correctly because human behavior is a

TABLE 1. Summary of literature review.

Ref.No	Signals	Features	Classifiers	Emotions	Accuracy in Per
[19]	EMG, ECG, RSP	Statistical, energy, entropy	LDA	Joy, Anger, Sadness, Pleasure	95
[20]	GSR, HP	No specific features stated	KNN, DFA	Sadness, Anger, Fear, Surprise, Frustration, Amusement	91.7
[21]	EMG, EDA, BVP, ECG, RSP	Running Mean Running Standard Deviation Slope	NN	Arousal, Valance	96.58
[22]	EDA, HR	No Specific Features Stated	Fuzzy Logic	Stress	99.5
[23]	BVP, EMG, EDA, RSP	Statistical Features	SVM, Fisher LDA	Amusement, Contentment, Disgust, Fear, Sad, Neutral	92
[24]	EMG, EDA, ECG, BVP, ST, RSP, Speech	Statistical Features, BRV, Zero- Crossing, MFCCs	KNN	Arousal, Valance	92
[25]	EDA, HR, EMG	No Specific Features	HMM	Arousal, Valance	81
[26]	EMG, EDA, ECG	No Specific Features	No Specific Classifier	Arousal, Valence	NA
[27]	EEG	No Specific Features	LSTM	Arousal, Valence and Liking	87
[28]	EEG	Min, Max Peak, Power	Feed Forward Neural Network	Stress, Normal	60
[29]	EEG	Statistical Features	SVM, ANN	Positive, Negative , Neutral	91
[30]	EEG, EMG, Temp, GSR, RSP	Different Features	MESAE	Arousal, Valence	77
[31]	EEG	ICA	MEMD, ANN, KNN	Arousal, Valence	82

very complex function and so many factors can effect human behavior based on emotions such as, their home environment or personal problems. Therefore there is dire need to identify and investigate other parameters, firstly that can help in improving efficiency of employees and secondly we can devise a way to determine and assess them more accurately using different techniques and methods.

One such parameter is working aptitude, as we know that it is an inborn skill of a person. In smart offices, in providing more conducive environment aptitude can play a vital role and will help in enhancing the efficiency of an employee, by assigning them tasks in light of their aptitude. Which will eventually lead to uplifting the performance of the whole office or organization. As aligned to these motivational examples, in this paper we propose a novel system; 'Multi Modal Aptitude Detection System (MMADS)' for smart offices. This monitoring system in smart offices can help in determining working aptitude of an employee that can be used by an employer or supervisor to bring improvements in employee's performance.

For example when employer or supervisor will be able to learn aptitude of his employee, he can better understand his job preferences, that is which particular tasks an employee can perform more better, more comfortably and thus can make them more beneficial for the office. MMADS in smart office, will also provide help in identifying the deficiencies present in terms of human resource for a particular office environment. Similarly MMADS can also be adopted to support employees, in providing them future directions for their successful careers in life.

MMADS in its initial phase; performs analytical skills detection, which is one dimension of aptitude. Whereas requirements for other dimensions of aptitude, such as IQ level, working interests, etc may also be investigated. This system helps in deciding, whether an employee is able to perform analytical tasks or not.

Contribution: This paper provides MMADS detailed framework, implementation pipeline and achieved outcomes. Initial prototype is based on physiological signals that includes Galvanic Skin Response (GSR), Body Temperature (Temp), Respiration (RSP) and Pulse Rate (HR). Implemented prototype is tested and validated using a 10 fold cross validation technique and achieve accuracy up to 96% using Convolutional Neural Networks (CNN). For further verification of results F1 scores are calculated that accumulates up to 0.906.

Paper Layout: Rest of the paper is organized as follows: Section II provides the literature review related to emotion detection and recognition systems as it is very close in working to our system, Section III discusses the detail framework of the system, Section IV elaborates the implementation details, Section V provides the overview related to the results, and Section VI concludes the paper.

II. LITERATURE REVIEW

In this section, a brief overview related to the research work done, so far, in the area of emotion recognition involving physiological signals is provided. In our daily life, emotions play a very important role in making decisions, planning, reasoning and various other activities [17].

Human emotion is a fine-grained behavior that emerges in specific situation. Learning emotions and taking actions depending on these emotions can produce more useful avenues. For example, in clinical empathy, in the treatment of patients especially those who cannot express themselves due to any disability, learning their actual emotional state can help doctors to treat them more effectively.

Emotion learning and recognition is becoming very important especially in brain computer interaction (BCI) applications and smart offices.

In smart offices after learning the real emotional state of an employee, an employer can treat them more effectively and fairly. Several studies have been performed on emotion recognition using physiological signals. Table 1 summarizes the work of different researchers published previously and provides details related to physiological signals captured, features extracted, classifiers used, emotions classified and achieved accuracy.

As can be seen, maximum accuracy of 99.5% for single emotion 'stress' is achieved in [22] using Electro Dermal Activity (EDA) and Heart Rate (HR) like physiological signals by using a fuzzy logic classifier. For two emotions maximum accuracy achieved is 96.58% in [21]. Authors have used Electromyography (EMG), EDA, Blood Volume Pressure (BVP), Electrocardiography (ECG) and Respiration (RSP) physiological signals and Neural Network (NN) classifier for classification. In [19], using four emotions maximum accuracy achieved is 95%. Authors have used EMG, ECG and RSP physiological signals with Linear Discriminant Analysis (LDA). Thejaswini *et al.* [29] have claimed to achieve an accuracy of 91% using electroencephalogram (EEG) physiological signal from brain. They have

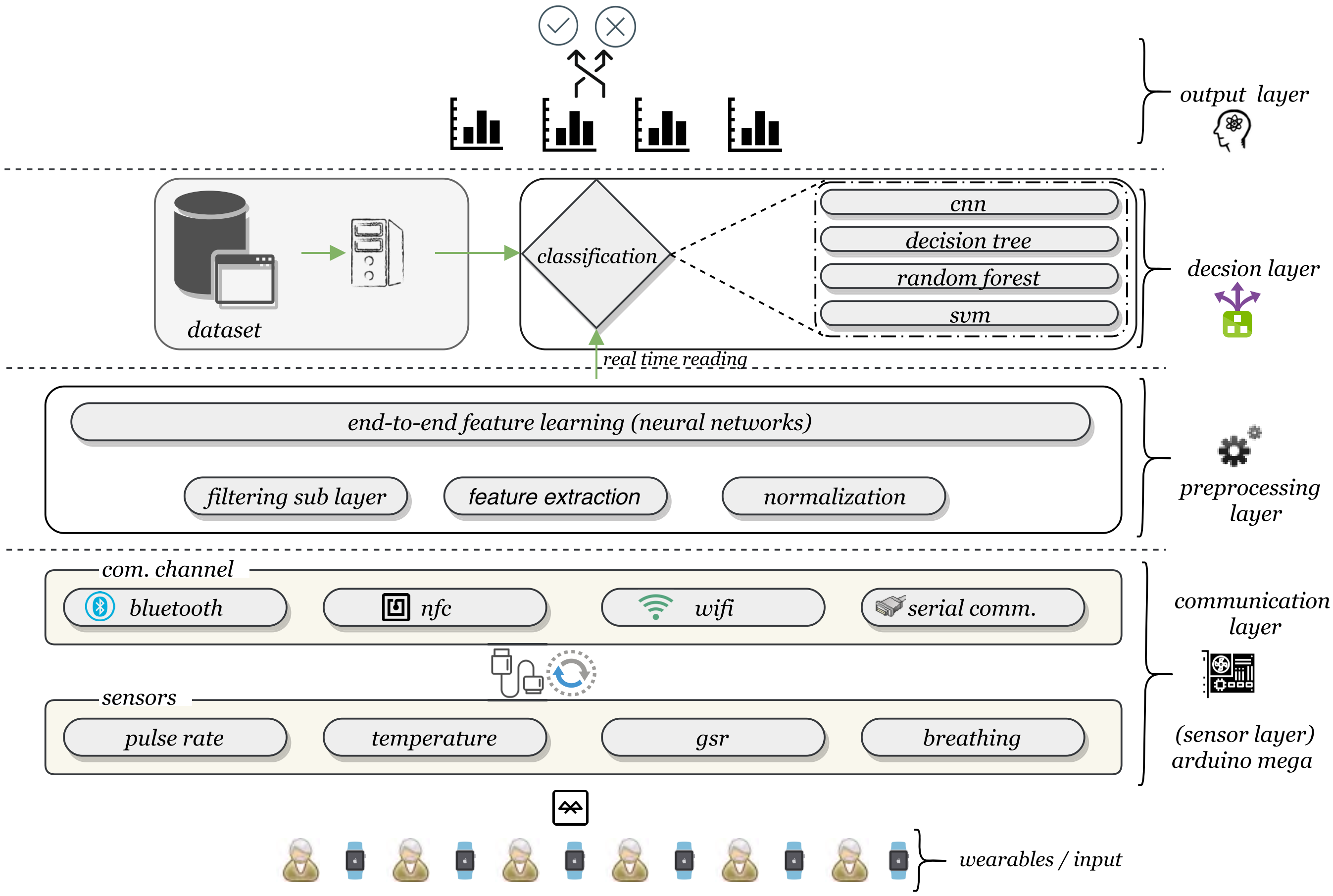


FIGURE 1. **Detailed layer framework of proposed system.**

used Support Vector Machine (SVM) and Artificial Neural Network (ANN) algorithms. Using ANN algorithm 91% accuracy has been achieved for emotions to be classified as positive, negative and neutral. Yin *et al.* [30] have used a combination of EEG, EMG, Body Temperature (Temp), Galvanic Skin Response (GSR), RSP physiological signals with an accuracy of 77%. For detecting two basic emotions arousal and valence, multiple-fusion-layer based ensemble classifier of SAE (MESAE) is being used. These figurative results provide an encouraging ground to march the research in this field forward.

Aptitude, as discussed earlier, is another avenue, which could help in improving workplace experience as it is an inborn skill of a person. We draw the inference that, to be able to learn the aptitude of an employee and assign them tasks accordingly will definitely help in dealing the employees effectively. And hence, smart offices suffice for the learning-based environments as necessary condition. Effective use of inborn skills (aptitude) will not only bring enhancement to the performance of the employees but will also benefit the employer.

The following section introduces our framework for the proposed system in detail, our proposed system helps in determining the working aptitude of an employee as a part of smart offices.

III. MULTI MODAL APTITUDE DETECTION SYSTEM's FRAMEWORK

The proposed framework comprises of five layers that are the sensor, communication, preprocessing, dataset, decision and output layers as shown in Figure 1.

Sensor Layer Numerous sensors are used as a main constituting part and will consist of a complete set of hardware that contains different sensor modules incorporated on a microcontroller called sensor unit. Sensor unit that works at this layer will be a complete set of sensors that generate sensed data in raw form related to a particular phenomenon and communicate it to the base station.

Communication Layer sensor unit at the sensor layer senses some phenomena that will need to be transferred from the sensor unit to base station where rest of the processing is performed with the sensed data. So communication layer is required and will serve as a gateway between the sensor unit and the base station. Communication layer contains communication modules such as blue tooth, NFC, Wi-Fi, serial communication etc. This layer can be implemented using an ad-hoc network or structured wired network however structured wired networks are recommended in the situation where privacy and security of data have high importance.

Pre Processing Layer when sensed data from sensor unit reach the base station through communication layer this data will be in the raw form which needs to be processed further in preprocessing layer. This layer will convert sensed raw data in a form that is required by decision layer to take decision accurately and reliably. The main task of this layer is to perform conversion of sensed data into a signal, filtering a signal to remove noises, extraction of the required data from the signal, normalization of data or any other operations necessary to make sensed data more reliable and work efficiently.

Dataset Layer for the different phenomenon to be identified with clarity, data related to a particular phenomenon is required and stored in properly labeled files called data set. This layer contains dataset files containing aptitude related data of employees which will be gathered as a result of experimental activity, later on, this dataset will be used to train the models. Normally, dataset files contain training data in a preprocessed or raw form that will be gathered from different subjects based on some experimental activity.

Decision Layer this layer will take final decision related to aptitude of an employee. The preprocessed test data from preprocessing layer and preprocessed training data from dataset layer will be used by applying a machine learning algorithm to take a decision regarding aptitude of a subject.

Output Layer this layer will receive a final decision from the decision layer and will apply further processing, based upon the required and achieved results. Finally, the output layer will use received results from the decision layer and further analyze to derive more concrete and refined outputs. Similarly, results received from different modalities are also analyzed here to come to a more refined conclusion.

IV. IMPLEMENTATION DETAILS

The design of our proposed framework was explained in detail in section III and how this design is implemented is being discussed in the following section.

A. SENSOR LAYER

In this layer, we have selected four wearable biomedical sensor modules to capture the true picture of vital body organs and these are pulse rate sensor module for detecting changes in heart functions, body temperature sensor module for detecting variations in body temperature, galvanic skin response sensor module for detecting variations in skin conductance and breathing sensor module for detecting variations in breathing patterns of a subject. As these are vital organs whose response changes to a particular condition or state of a body [17]. These sensor modules are mounted and configured on Arduino-Uno-Mega micro-controller board, a schematic diagram of sensor modules connection is depicted in Figure 2.

Sensor modules are integrated and image that controls them is burned in the ROM of micro-controller to channelize data values generated by these sensors as depicted in Figure 3.

B. COMMUNICATION LAYER

For communicating sensed data values received by micro-controller a blue tooth module is mounted, this blue tooth module communicate sensed values to a base station after establishing a connection with the base station. For a long-range communication, this blue tooth module can be replaced with a Wi-Fi module. The serial COM3 port is used to read received sensed data at the base station as depicted in Figure 4.

This sensed data received through serial communication port at the base station is stored in a CSV format.

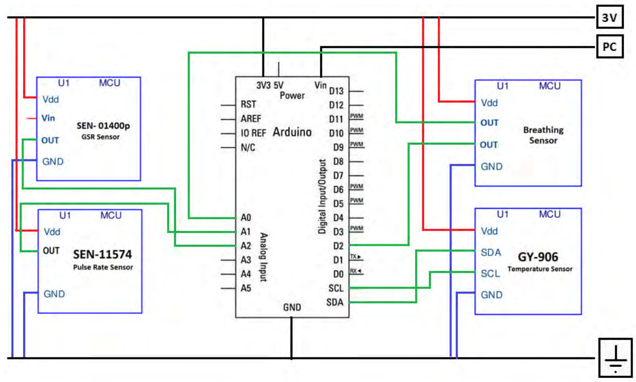


FIGURE 2. Schematic diagram of interconnected sensor modules.

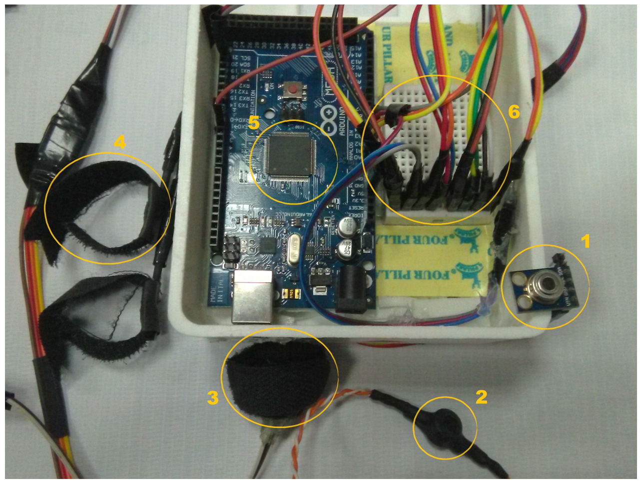


FIGURE 3. Sensors interconnected with arduino micro-controller (different parts of hardware used are labeled in numbers whereas 1 show infrared based body temperature sensor, 2 shows breathing sensor, 3 is pulse rate sensor, 4 is GSR sensor, 5 depicts Arduino-Uno-Mega micro-controller and 6 shows interconnection wires).

C. PRE PROCESSING LAYER

This layer performs all processing needed to make a signal ready for decision making. First of all sensed data from CSV file is converted to a signal. Figure 5(B) depicts all four modalities as signals in raw form. As can be seen in Figure 5(B) against GSR raw form of signal, it contain huge number of spikes as noises whereas most of signal values lies between 250 to 500 domain.

To make these signals ready for decision making, signals need to be passed through normalization and noise removal phase. In the normalization phase of preprocessing, operations of mapping to the normal domain of individual

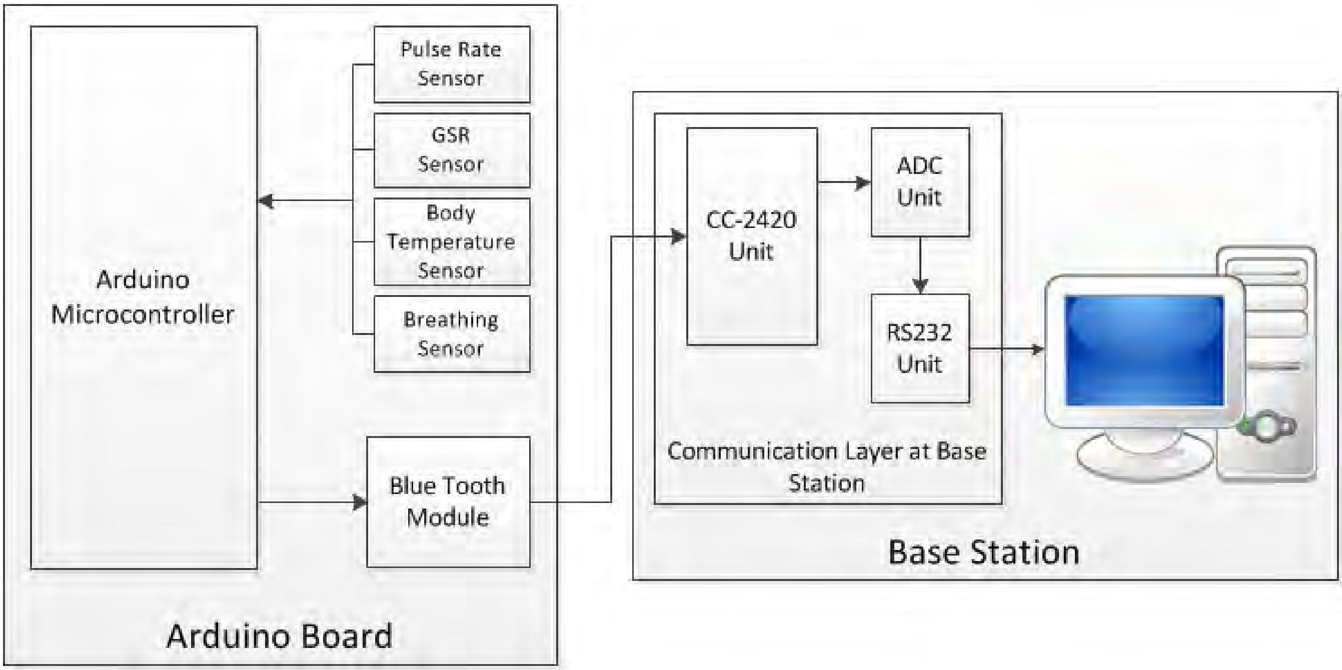


FIGURE 4. **Communication layer hierarchy.**

modalities where necessary is performed after that signal becomes normalized. For an instance signal of GSR, breathing and temperature need not be normalizing, they only contain noises. Whereas pulse rate signals need normalizing the data values. After normalizing, Figure 5(A) depicts a normalized form of pulse rate signal. For mapping to normalization domain, first mean of the raw signal is calculated which is then compared with the normal range of domain. Based on comparison a pivot value *'P'* is derived. This value is then utilized in the normalization process. Normalization process is depicted in equation 1 below.

$$S(t) = \begin{cases} S(t) & \text{if } 40 <= S(t) <= 140 \\ S(t) + P & \text{if } S(t) < 40 \\ S(t) - P & otherwise \end{cases} \tag{1}$$

where S(t) is signal's sample value at time interval t and P is pivot value.

After normalization phase signals are then passed through noise removal phase, where mean of data values are calculated on windows of small size. Size of the window is kept so small that taking mean will remove the noises from the signal without disturbing the overall distribution of the signal. As can be seen in Figure 5(A) the distribution of both normalized and preprocessed signals are alike. Similarly, if we compare the two signals in Figure 5(A) it can be established that the number of instances in signal is reduced after noise removal operation almost the ten times without affecting the distribution of the signals. Means and standard deviation of both normalized and preprocessed signals are calculated which are given in Table 2. After noise removal the signal becomes preprocessed signal. Mathematical model of noise removal operation is shown in equation 2, whereas Figure 5(A) shows the normalized and preprocessed signal.

$$\mu = (\sum_{t=1}^{N} S(t)) \backslash N \tag{2}$$

In equation 2 μ is averaged of N samples and S(t) is sample value at time interval t. The results obtained from the function described in equation 2 are served as input to our proposed CNN model in sub section 'Decision Layer'.

Comparison of all four signals from the four modalities can be made, Figure 5(B) shows a raw form of a signal from all the four modalities whereas Figure 5(C) depicts a preprocessed form of signals. Now it can be seen that most of the noise spikes and outliers are removed from the signals with enhancement in the signal's required distribution.

After preprocessing data is ready to feed to decision layer to derive conclusions, operations performed in the decision layer are discussed in coming sections.

D. DATASET LAYER

This layer contains data files related to the ground truth called dataset, collected after carefully designing and implementing a proper experimental environment. As the proposed system is an indigenous system so no related data set was available therefore a new dataset was created in order to have a proper ground truth. For the sake of simplicity and to narrow down the scope, the initial prototype of the proposed system has only two classes (have analytical skills or no analytical skills) in future classes can be increased like no, ordinary,

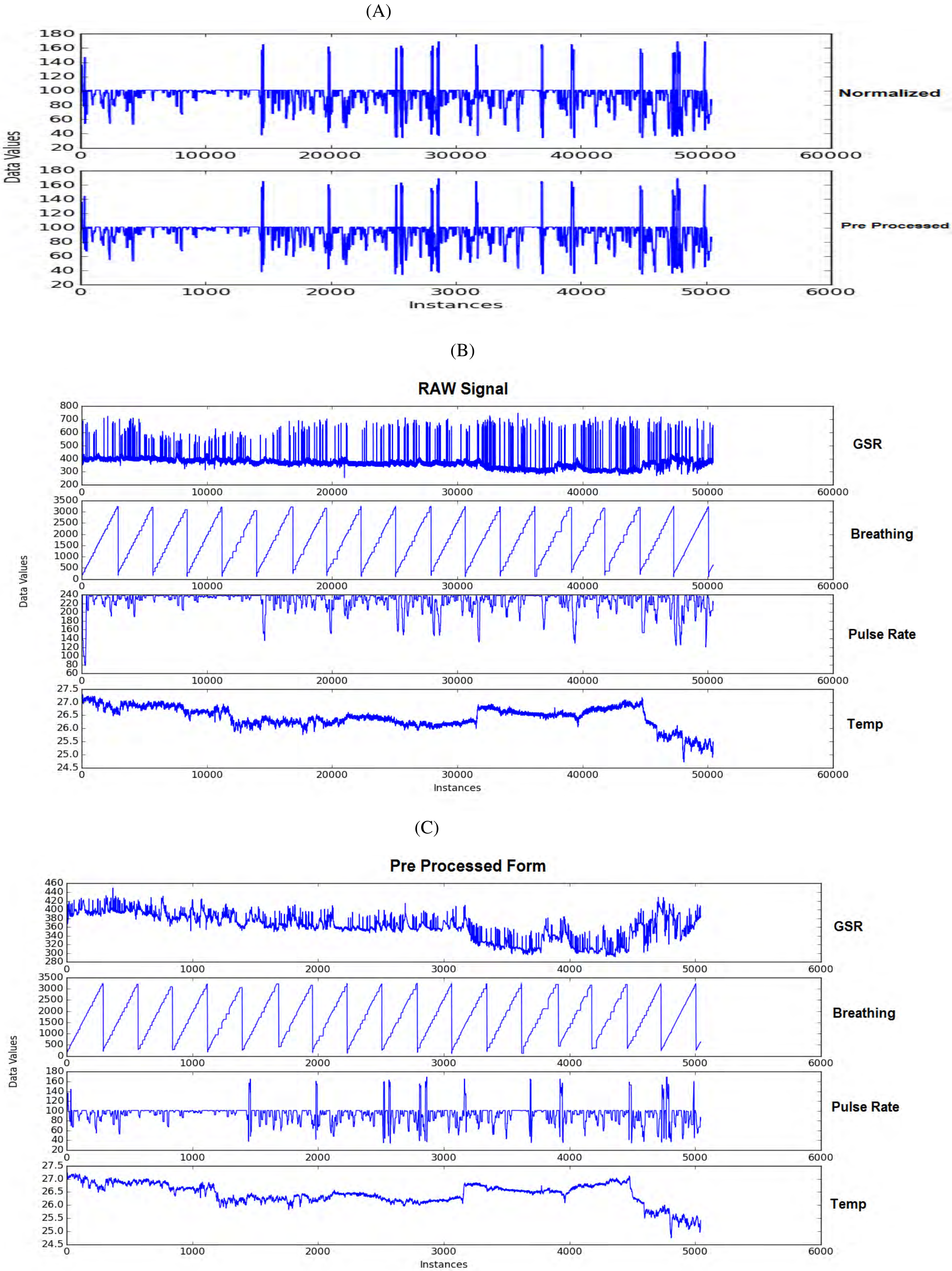


FIGURE 5. Comparison of signals (A) Shows comparison of normalized and pre-processed form of pulse rate signal. (B) Signals in raw form from all 4 modalities. (C) Pre-processed form of signals from all 4 modalities.

moderate, good, better and extraordinary analytical capabilities. An experiment is designed in a way that initial prototype of a system is deployed that contain four wearable sensors mounted on an Arduino micro-controller, details are already discussed in above subsections, Figure 6 presents data flow for collecting ground truth.

TABLE 2. Comparison of raw and preprocessed form of pulse rate signal.

Signal	Mean	Standard Deviation
Normalized Signal (Figure 5(A))	93.2	17.2
Preprocessed(Figure 5(A))	93.2	16.81

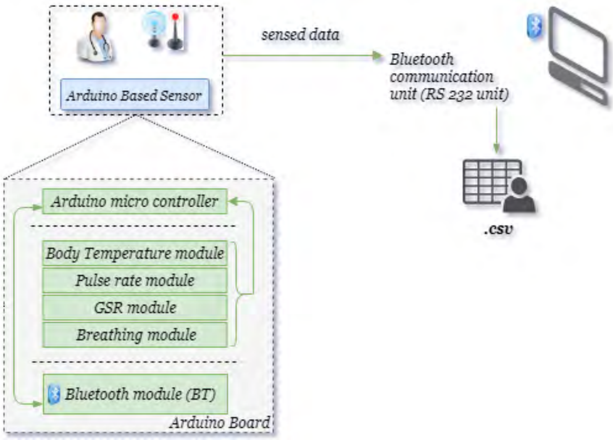


FIGURE 6. Ground truth collection.

After deploying the system main goal was to collect the state of the four modalities (heart, skin, lungs, and temperature) at a time when a person is performing an analytical task. To comply with an analytical reasoning MCQs based test has been designed, test questions were taken from analytical reasoning portion of the International GRE [32]. A subject is given the test to solve for some fixed time interval after attaching the wearable sensors.

This test contains different scenarios and analytical reasoning questions and answers in order to provoke the analytical capabilities of a subject. While the subject is solving the test, wearable sensors are collecting state of different vital organs of the body of a subject. After solving the test are marked and cumulative scores are calculated which are then matched against the minimum threshold value, if it is less or equal to threshold value then this collected data is labeled as "no analytical skills" otherwise it is labeled as "have analytical skills". After collecting the data it is then stored in the proper files to make dataset. The experiment was performed in a controlled environment. Proper office setup was created with a table, desktop computer, sitting chair, and other necessary equipment. Other factors such as room temperature, humidity, and noise were also taken into consideration. Ground truth is collected from 20 different individuals, both male and female of ages from 22 to 45 years, which was then stored in proper CSV files as raw data. This data will use to train our machine learning models, before training these files were passed through the preprocessing phase for normalization and removal of noises and finally this data will use in training.

E. DECISION LAYER

In this layer, different models are trained using collected dataset for decision making. Trained models then used to classify raw signal captured from an employee in real office environment working on assigned tasks. In this study different models like KNN, SVM, Decision Tree, Naive Bayes, Random Forest, and CNN are implemented and validated for checking the accuracy with real-time data. For implementing the (KNN, SVM, Decision Tree, Naive Bayes and Random Forest) the preprocessed signals from preprocessing phase is then passed through the smoothing phase, smoothed signal is then transformed from time domain to frequency domain using a Fast Fourier Transform method, then a feature vector based on different features is generated from the signal in frequency domain. This feature vector is then used to train these models. Features and their description are given in Table 3 below.

TABLE 3. Features description.

Feature	Description
Minimum	Minimum Frequency of a Signal
Maximum	Maximum Frequency of a Signal
Mean	Mean of Overall Frequency Signal
Standard Deviation	Signal Deviation From the Standard

As there are four signals from four different modalities and calculating four features per signal therefore we have 16 features extracted from one whole file and inserted as a single row in feature vector. This features based vector now become ready to train our models.

Convolutional Neural Network (CNN) model is built using *convolutional layer with RELU activation, pooling, dropout, flatten, fully connected layer with soft-max activation.* Order of the layers with detail description is provided in Figure 7 where as details of the layers are provided below.

Convolutional Layer: This layer use filter to perform convolution operation over the input signals. Filter is a two dimensional array which slides through two dimensional signal's array from all four modalities and convolves it. Another parameter stride controls the convolutional operation and decides how much signal should be convolved with the filter. Output of this operation is a feature map that is stored in the form of vector and contains learned features from signal. Equation 3 below provides the detail operation of convolution.

$$Y_k = \sum_{n=1}^{N-1} X_n H_{k-n} \tag{3}$$

where X is signal, H is filter, and N is the number of elements in X. The output vector is Y, a feature map vector. The subscripts denote the kth element of the vector.

Rectified Linear Activation Unit (RELU): Convolutional operation is a linear function so after every linear operation a non linear activation function is used. This activation function introduce non linearity and sparsity to the network due to which the network becomes more robust to small changes such as noise in the input. The rectifier linear unit is a well known activation function for deep learning.

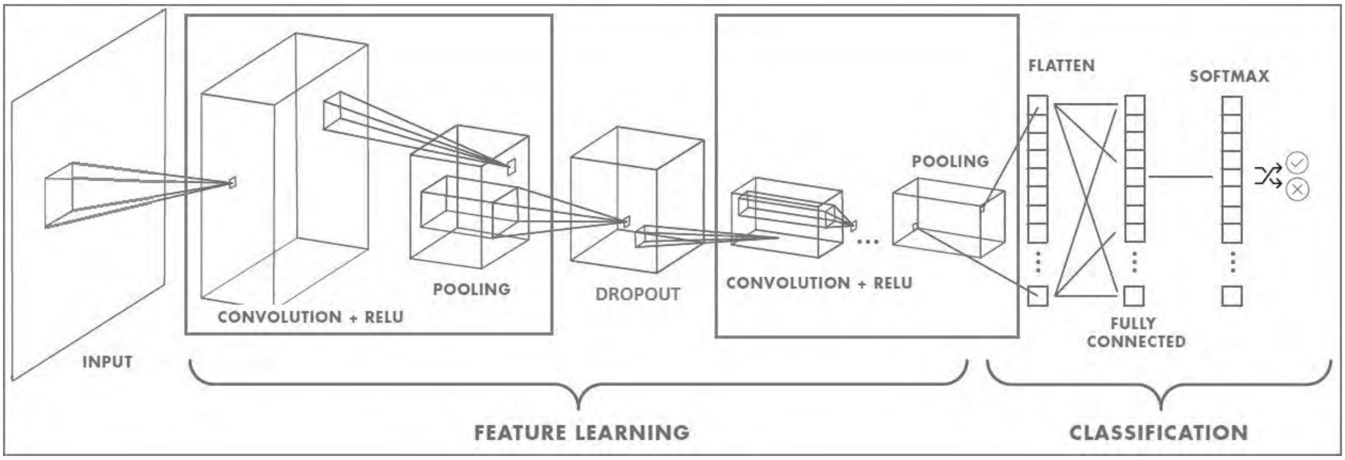


FIGURE 7. CNN detail description.

Equation 4 shows the RELU function.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \tag{4}$$

Pooling: For performing the pooling function max-pooling operation is applied. In this technique a window of arbitrary size is selected, and windows can be overlapping. This window is applied to the patch selected from feature map and maximum value is selected from the resultant patch to represent it in output feature map generated by this layer. Operation of this layer is depicted in equation 5.

$$B_j = \max(A_i^{n*n} U(n, n)) \tag{5}$$

In equation 5 B_j is the jth element of output feature map vector B, A_i^{n*n} is n x n patch from the input feature map produced from previous layer and $U(n, n)$ is the window of size n x n.

Drop Out: This layer is used to randomly drop out the neurons in order to handle over-fitting problems in Neural Networks. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass.

Flatten: This layer serves as an input to fully connected layer. This layer flattens and reshapes the convolved data from the previous layers, into the form that is required by the fully connected layer to classify learns features.

Fully Connected: This is the final layer of our CNN architecture where final classification is carried out related to aptitude of a subject. The neuron in the fully-connected layer detects a certain feature, for instance a particular sequence of spikes in a signal. It retains and stores its values and verify against ground truth and label them as either 'Have Analytical Skills' or 'No Analytical Skills'.

Softmax: This function uses probability to compute distribution of the k output classes. Hence, in our network it is a last layer. Soft-max function finally decides that an input signal belongs to which class after computing the probability distribution for all k classes. Soft-max function is depicted in equation 6.

$$P_j = \frac{e^{xj}}{\sum_1^k e^{xk}} \quad for\ j = 1, \ldots.k \tag{6}$$

where x is the net input. Output values of p are between 0 and 1 and their sum equals to 1.

All these models are implemented using keras and tensor flow in python.

F. OUTPUT LAYER

Main contribution of this layer is to finalize output of the system, in fact proposed system receives body state from the four different modalities as already discussed, now the main question that arises here is whether individual state of organs is used to come to a final conclusion or after combining the states from all the modalities, should represent the final body state. The proposed system is using the second approach. First, we receive the individual state of the organs and then they are combined to represent a state of body which is then used to derive a final conclusion.

V. RESULTS

Evaluation Criteria We have evaluated our models' performance by using the following criteria which is based on following assumptions. The signal labeled as ''have analytical skills'' is positive signal and signal that is labeled as ''no analytical skills'' is a negative signal.

True Positive (TP): It is number of total positive signals classified correctly by the model. *False Negative (FN):* It is number of total positive signals classified incorrectly by the model. *False Positive (FP):* It is number of total negative signals classified incorrectly by the model. *True Negative (TN):* It is number of total negative signals classified correctly by the model.

Specificity: Specificity is also called False positive rate and is calculated as the number of false positive predictions divided by the total number of negatives. Total number of

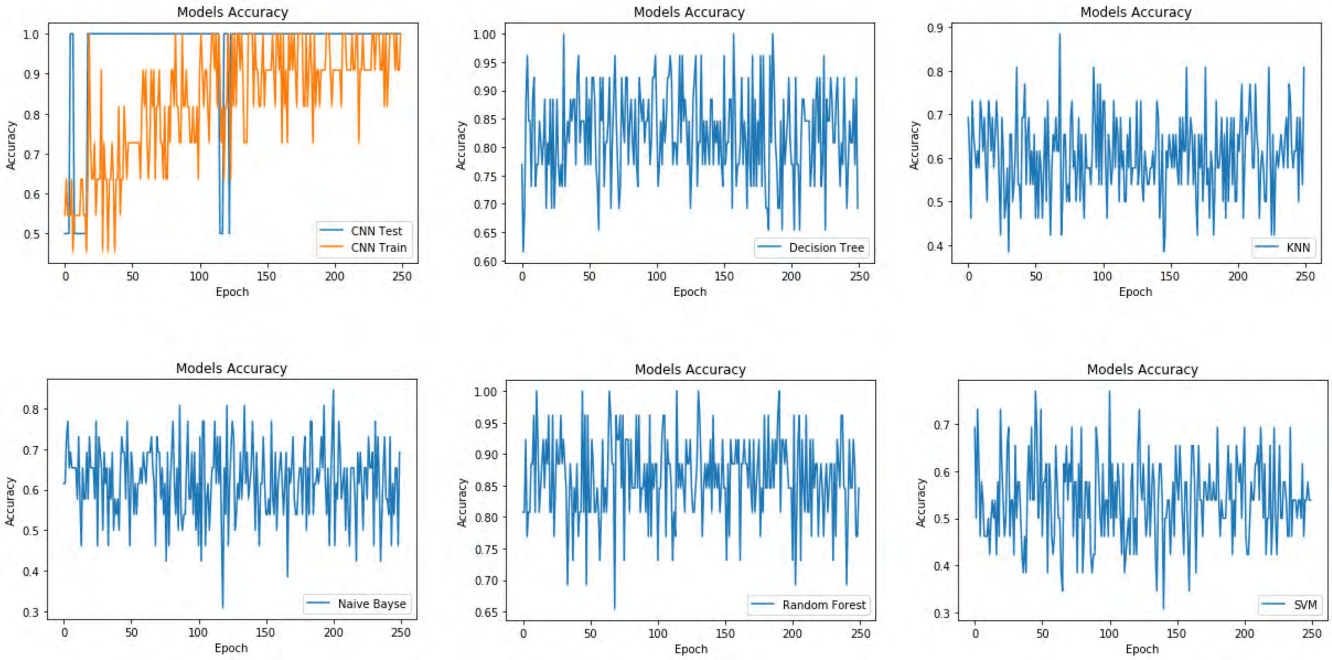


FIGURE 8. Accuracy results of all six models.

TABLE 4. Detail comparison of models accuracy in percentages.

S.No	Model	Epochs	Maximum Accuracy	Minimum Accuracy	Average Accuracy
1	CNN (Training)	250	100	45	83
2	CNN (Validation)	250	100	50	96.4
3	Decision Tree	250	100	61.5	82.4
4	KNN	250	88.4	38.5	60.4
5	Naive Bayse	250	84.6	30.8	62.1
6	Random Forest	250	100	65.4	86.3
7	SVM	250	76.9	30.8	53.9

negative signals is sum of negative signals predicted as negatives(TN) and Number of negative signals predicted falsely as positive(FP) it is calculated as:

$$Specificity = \frac{FP}{TN + FP}$$

Sensitivity: Sensitivity or recall is also called true positive rate and is calculated as the number of correct positive predictions divided by the total number of positives. Total number of positive signals are signals that are predicted correctly(TP) plus the positive signals predicted incorrectly(FN) it is calculated as:

$$Sensitivity = \frac{TP}{TP + FN}$$

Precision: Precision is calculated as the number of correct positive predictions divided by the total number of positive predictions it is calculated as.

$$Precision = \frac{TP}{TP + FP}$$

F1-Score: F1-score is a harmonic mean of precision and recall or F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. When we have un even class distribution then this metric is more useful than accuracy and it is calculated as.

$$F1 - Score = \frac{2 * Prec * Sens}{Prec + Sens}$$

For validation purpose, all the models are trained and tested using a 10-fold cross-validation method. In this method, all the data is divided into ten equal size chunks where nine out of ten chunks are used to train the model and remaining one is used to validate the accuracy of the model [33]–[36]. Figure 8 shows the accuracy obtained from all the six models.

As shown in the Figure 8 CNN is producing the best results as for most of the epochs accuracy lies between 90 to 100%, after that random forest is achieving good accuracy results as for most of its epochs its lies between 80 to 95% and on the third place we have decision tree model as its accuracy lies between 80 to 90%. Table 4 shows the detail accuracy comparison of all the six models.

For the verification of the models confusion matrix and f1 scores are also generated as shown in Table 5

TABLE 5. Confusion matrix, F1 scores, sensitivity, specificity and recall.

Models	True Positive	False Negative	False Positive	True Negative	F1 Score	Sensitivity	Specificity	Recall
CNN	126	14	12	108	**0.906**	**0.90**	**0.90**	**0.91**
RF	120	8	22	110	0.884	0.94	0.83	0.85
DT	126	26	10	98	0.87	0.83	0.91	0.93
KNN	111	46	51	52	0.626	0.71	0.50	0.69
NB	54	82	12	112	0.614	0.40	0.90	0.82
SVM	148	0	112	0	0.419	1.00	0.00	0.57

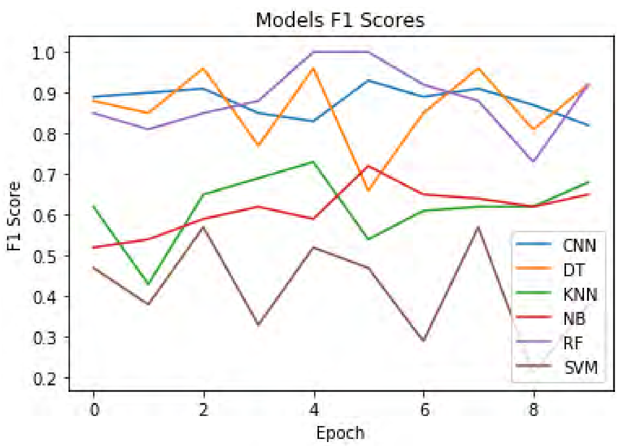


FIGURE 9. F1-Score of all six models.

and detail comparison of F1-Scores of all the models is depicted in Figure 9, validation process is performed on total 260 samples.

As can be seen in Table 5 best accuracy results are justified by the CNN model as its true positive are 126 out of 140 and true negative are 108 out of 120. It means that from 140 positive samples it has miss-classified only 14 samples and vice versa. Similarly, the results for other models can also be interpreted, as can be seen, that SVM model has miss classified all the negative class samples that are "no analytical skills" and perfectly classified all the positive class samples that are "have analytical skills". The reason for this behavior is: the model cannot mark the clear distinction line between the two classes, it has intermixed badly the two classes and it is because of the nature of the data that is making it behave poorly. Actually, data won't contain the sparsity required for SVM to behave in a better way as the samples have not much difference in terms of nonlinearity belonging to two different classes. Same is the reason for KNN and NB poor accuracy.

VI. CONCLUSIONS AND FUTURE WORK

This paper proposes a detailed framework, provides detail implementation pipeline to make it a reality and share results for a novel system named "Multi Modal Aptitude Detection System". This system is proposed and investigated with the intention that it will help in smart offices in providing a more conducive environment to employees working there. If smart offices are being made self-sufficient in investigating and providing more conducive environment to their employees, after learning their aptitude it will contribute a lot towards achieving main goals of smart offices. However, proposed system has only focused at one dimension of the aptitude that is analytical skills of a subject. Initially, this system is implemented and tested for a controlled office environment to be used as a part of smart offices and will help the employers in determining analytical aptitude of employees that is whether an employee can perform analytical tasks or not. Results obtained from the implementation of the initial prototypes are, CNN provides the accuracy of 95% whereas detail comparison of all the models is provided in Table 4. From the result obtained, the proposed system can be recommended to be used in smart offices for the well-being of both the employers and employees.

As a future work, the system can be improved to incorporate detection of moderate, good, better, and exceptional analytical capabilities. Apart from it detection of capabilities related to IQ, working aptitude etc. can be incorporated in the system to make it more generalized. Currently system is implemented from smart offices perspective later on its scope can be broadened to smart homes and smart cities, in smart homes parents can be assisted in deciding more precisely about the future career lives of their kins and in smart cities system can assist higher officials in better utilizing the potential present as human resource in their cities after determining their aptitude.

REFERENCES

[1] Y. Li *et al.*, "Multimodal BCIs: Target detection, multidimensional control, and awareness evaluation in patients with disorder of consciousness," *Proc. IEEE*, vol. 104, no. 2, pp. 332–352, Feb. 2016.

[2] S. Prasanna and S. Rao, "An overview of wireless sensor networks applications and security," *Int. J. Soft Comput. Eng.*, vol. 2, no. 2, pp. 538–540, May 2012.

[3] T. F. Quatieri *et al.*, "Using EEG to discriminate cognitive workload and performance based on neural activation and connectivity," MIT Lincoln Lab., Lexington, MA, USA, Tech. Rep., 2016.

[4] N. Sazak and H. Abdullah, "The importance of using wireless sensor networks for forest fire sensing and detection in Turkey," in *Proc. 5th Int. Adv. Technol. Symp. (IATS)*, May 2009, pp. 13–15.

[5] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "Spins: Security protocols for sensor networks," *Wireless Netw.*, vol. 8, no. 5, pp. 521–534, Sep. 2002.

[6] F. Stajano and R. Anderson, "The resurrecting duckling: Security issues for ubiquitous computing," *Computer*, vol. 35, no. 4, pp. 22–26, Apr. 2002.

[7] T. Ojha, S. Misra, and N. S. Raghuwansh, "Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges," *Comput. Electron. Agricult.*, vol. 118, pp. 66–84, Oct. 2015.

[8] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2688–2710, Oct. 2010.

[9] B.-S. Lin, A. M. Wong, and K. C. Tseng, "Community-based ECG monitoring system for patients with cardiovascular diseases," *J. Med. Syst.*, vol. 40, no. 4, p. 80, Apr. 2016.

[10] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using Android smartphones with accelerometers," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. Workshops (DCOSS)*, Jun. 2011, pp. 1–6.

[11] N. Batista, R. Melício, J. C. O. Matias, and JPS Catalão, "Photovoltaic and wind energy systems monitoring and building/home energy management using ZigBee devices within a smart grid," *Energy*, vol. 49, pp. 306–315, Jan. 2013.

[12] N. Maisonneuve, M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels, "Citizen noise pollution monitoring," in *Proc. 10th Annu. Int. Conf. Digit. Government Res., Social Netw., Making Connections Between Citizens, Data Government*, May 2009, pp. 96–103.

[13] R. Tan, G. Xing, J. Chen, W.-Z. Song, and R. Huang, "Quality-driven volcanic earthquake detection using wireless sensor networks," in *Proc. 31st IEEE Real-Time Syst. Symp. (RTSS)*, Nov./Dec. 2010, pp. 271–280.

[14] H. Zhou and B. Goold, "A domestic adaptable infant monitoring system using wireless sensor networks," in *Proc. IEEE 34th Int. Perform. Comput. Commun. Conf. (IPCCC)*, Dec. 2015, pp. 1–2.

[15] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *BioMed Res. Int.*, vol. 2017, Aug. 2017, Art. no. 8317357.

[16] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Emotion*, vol. 8, no. 10, pp. 355–358, 2017.

[17] M. Ali, A. H. Mosa, F. A. Machot, and K. Kyamakya, "Emotion recognition involving physiological and speech signals: A comprehensive review," in *Recent Advances in Nonlinear Dynamics and Synchronization*. Cham, Switzerland: Springer, 2018, pp. 287–302.

[18] Z. Mohammadi, J. Frounchi, and M. Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Comput. Appl.*, vol. 28, no. 8, pp. 1985–1990, Aug. 2017.

[19] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[20] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP J. Adv. Signal Process.*, vol. 2004, p. 929414, Dec. 2004.

[21] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using biosensors: First steps towards an automatic system," in *Affective Dialogue Systems* (Lecture Notes in Computer Science), vol. 3068, E. André, L. Dybkjær, W. Minker, and P. Heisterkamp, Eds. Berlin, Germany: Springer, 2004, pp. 36–48.

[22] A. de Santos Sierra, C. S. Avila, J. G. Casanova, and G. B. del Pozo, "A stress-detection system based on physiological signals and fuzzy logic," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4857–4865, Oct. 2011.

[23] C. Maaoui and A. Pruski, "Emotion recognition through physiological signals for human-machine communication," in *Cutting Edge Robotics*. Rijeka, Croatia: InTech, 2010.

[24] J. Kim, "Bimodal emotion recognition using speech and physiological changes," in *Robust Speech Recognition and Understanding*. Rijeka, Croatia: InTech, 2007.

[25] D. Kulic and E. A. Croft, "Affective state estimation for human–robot interaction," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 991–1000, Oct. 2007.

[26] W. Yang, M. Rifqi, C. Marsala, and A. Pinna, "Physiological-based emotion detection and recognition in a video game context," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–8.

[27] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Emotion*, vol. 8, no. 10, 2017.

[28] R. Mahajan, "Emotion recognition via EEG using neural network classifier," in *Soft Computing: Theories and Applications*. Springer, 2018, pp. 429–438.

[29] S. Thejaswini, K. M. R. Kumar, S. Rupali, and V. Abijith, "EEG based emotion recognition using wavelets and neural networks classifier," in *Cognitive Science and Artificial Intelligence*. Singapore: Springer, 2018, pp. 101–112.

[30] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Comput. Methods Programs Biomed.*, vol. 140, pp. 93–110, Mar. 2017.

[31] A. Mert and A. Akan, "Emotion recognition from EEG signals by using multivariate empirical mode decomposition," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 81–89, Feb. 2018.

[32] W. B. Kernighan and D. M. Ritchie, *Barron's GRE*, 22nd ed. Hauppauge, NY, USA: Barrons Educational Series, 2017.

[33] P. Kumari, L. Mathew, and P. Syal, "Increasing trend of wearables and multimodal interface for human activity monitoring: A review," *Biosensors Bioelectron.*, vol. 90, pp. 298–307, Apr. 2017.

[34] M. Elgendi, A. Al-Ali, A. Mohamed, and R. Ward, "Improving remote health monitoring: A low-complexity ECG compression approach," *Diagnostics*, vol. 8, no. 1, p. E10, Jan. 2018.

[35] M. M. Umar, A. Mehmood, H. Song, and K.-K. R. Choo, "I-Marks: An iris code embedding system for ownership identification of multimedia content," *Comput. Elect. Eng.*, vol. 63, pp. 209–219, Oct. 2017.

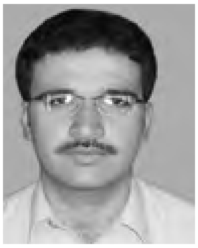
[36] Z. Zhang, A. Mehmood, L. Shu, Z. Huo, Y. Zhang, and M. Mukherjee, "A survey on fault diagnosis in wireless sensor networks," *IEEE Access*, vol. 6, pp. 11349–11364, 2018.

MOHAMMAD TEHSEEN is currently pursuing the Ph.D. degree with the Department of Computer Science, University of Peshawar, Pakistan, where he is also a Ph.D. Scholar. He has published several papers in reputed journals and conferences. His research interests include wireless sensor networks, middleware, the Internet of Things, wireless sensor-based applications, deep learning, and signal processing.

HUMA JAVED received the Ph.D. degree in software engineering from the School of Computing and Mathematical Sciences, Liverpool John Moores University, U.K., in 2008. She is currently an Assistant Professor with the Department of Computer Science, University of Peshawar, Pakistan. Her research interests include wireless sensor networks, middleware, the Internet of Things, and wireless sensor-based applications. She has served as a Reviewer for several journals.

AMJAD MEHMOOD received the Ph.D. degree in wireless networks from the Kohat University of Science and Technology (KUST), Kohat, in 2014. He held two Post-doctoral positions from the SONG Lab, University of Virginia, Charlottesville, VA, USA, and from the Guangdong Provincial Key Laboratory on Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, Maoming, China. He has got the training of Especially Assigned Researcher from the Queensland University's Data 61 Lab, The University of Hong Kong, and the University of Beijing, China. He is a Cisco Certified Trainer and Java Certified Programmer. As far as job experience is concerned, in 2003, he joined KUST, where he is currently serving as a Professor and the Coordinator of the MS/ PhD program. He is also the Director of the Research Lab at KUST, named as the Reliable Optimal Smart Scalable and Efficient Lab. He has published over 60 academic articles in peer-reviewed international journals and conferences around the world. He is interested to work in the areas of cyber–physical systems, the IoT, connected vehicles, wireless communications and networking, optical communications and networking, smart grid communications and networking, security issues in wireless networks, big data, cloud computing, and fault diagnosis in WSNs. He has visited many countries for presenting his research work like Australia, Hong Kong, China, South Korea, Malaysia, and Thailand. He has supervised many students of BCS, MCS, M.S., and Ph.D. in the above-mentioned interests. He has remained as a part of reviewing and organizing different workshops, seminar, and training sessions on different technologies. As far as his social experience is concerned, he is a Professional member of the ACM and the KUST University Branch Coordinator. Furthermore, he has served as a TPC member, Reviewer, and Demo Chair for numerous international conferences, including CCNC, SCPA, WICOM, INFOCOM, SCAN, and so on. He is currently elected as the Chair IEEE Member Development and Award Committee for the period of three years (until 2021). In addition, he has been serving as a Reviewer, Associate Editor, and Editor for many peer reviewed international journals.

MUHAMMAD AMIN is currently pursuing the Ph.D. degree with the Department of Computer Science, Institute of Management Sciences, Peshawar, Pakistan, where he is also a Ph.D. Scholar. He has published several papers in reputed journals and conferences. His research interests include image processing, malware detection, wireless sensor-based applications, deep learning, and signal processing.



ISHTIAQ HUSSAIN received the Ph.D. degree from the Department of Computer Science, University of Peshawar, in 2018. He is currently with the Higher Education Department, Pakistan. He has published several papers in reputed journals and conferences. His research interests include wireless sensor networks, process migration, the Internet of Things, virtual reality, and augmented reality.

BILAL JAN received the M.S. and Ph.D. degrees from the Department of Control and Computer Engineering, Politecnico di Torino, Italy, in 2010 and 2015, respectively. He is currently an Assistant Professor with the Department of Computer Science, FATA University, Kohat, Pakistan. He has published several papers in reputed journals and conferences. His research interests include wireless sensor networks, the Internet of Things, big data, deep learning, and high-performance computing. He has served as a Reviewer for several journals.