# Hierarchical deep neural network for mental stress state setection using IoT based biomarkers

Akshi Kumar[a,**], Kapil Sharma[b], Aditi Sharma[a]

[a]Department of CSE, Delhi Technological University, Delhi and 110042, India
[b]Department of IT, Delhi Technological University, Delhi and 110042, India

ABSTRACT

Stress is a psychological condition in which a person feels overwhelmed with pressure. It can be positive, keeping us alert and motivated or negative causing emotional and physical wear-tear. The body's autonomic nervous system has a built-in stress response that causes physiological changes to allow the body to combat stressful situations. Bio-signals are biomarkers depicting these physiological changes during chronically activated situations. Only trained medical practitioners can measure such indicators which can be tedious and time consuming, delaying early identification and timely intervention. With the availability of IoT based sensors for healthcare, these biomarkers can be tracked using various wearable devices. Motivated by the need to design a model for mental stress state detection using sensor-based bio-signals, this research proffers multi-level deep neural network with hierarchical learning capabilities of convolution neural networks. A multivariate time series data consisting of both wrist-based and chest-based sensor bio-signals is trained using a hierarchy of networks to generate high-level features for each bio-signal feature. A model-level fusion strategy is proposed to combine the high-level features into one unified representation and classify the stress states into three categories. A superlative performance accuracy of 87.7% is achieved using the proposed network, which outperforms the state-of-the-art results.

## 1. Introduction

The craving to succeed in this fast-paced life takes away the time to overhaul oneself. An individual encounters constant pressure to excel at everything, job stress, nagging by parents and peer pressure leading to increased risk of developing mental health problems. Undeniably, stress is a common problem in modern life psychology. However, mental health is not visible to anyone and it is hard to identify the real mental health status of a person. In contrast, physical health whenever deteriorates has observable signs & symptoms and we seek immediate diagnosis, advice and treatment for it from the healthcare professionals. Moreover, some of the mental health conditions like depression, acute stress syndrome, anxiety and insomnia have common symptoms, so classifying the correct type becomes challenging. For example, stress and anxiety both have symptoms like a dry mouth, sudden sweating, and increased breathing rate. Therefore to classify them as stress or anxiety, the subject needs to be evaluated for an extended period as anxiety attack lasts only for few minutes and is often triggered by an external stimuli. The physical and psychological impacts can be cyclically linked: emotional distress and poor mental health can trigger or flare a physical health problem and, as a result, cause further distress. Likewise, poor physical health can lead to an increased risk of developing mental health problems. A mild amount of stress can be favorable, as it has been observed that a person gives near-optimal works performance under mild-stress. Eustress or beneficial stress [1] is often related to a positive challenge as compared to distress which has negative implications. However, prolonged and chronic stress can severely impact person's health, affect the whole body and increase the risk of developing certain illnesses. It can have several physical or psychological symptoms, which can make functioning on a daily basis more challenging.

Formally, mental illnesses are health conditions involving changes in emotion, thinking or behavior (or a combination of these) [2]. The general cognitive function is hindered to an extent that it can trigger inappropriate responses because those responses are based upon inaccurate thoughts. That is, the person finds difficult to stay focused, process information, store it in memory, and accurately respond. Mental illness is conceptualized as a clinically significant behavioral dysfunction or psychological syndrome. There many different categories of mental/ psychological disorders defined in the ICD-10, 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD), a medical classification list by the World Health Organization (WHO) known as mental and behavioral disorders, ICD codes F00 to F99.

The healthcare industry has radically changed as the Internet

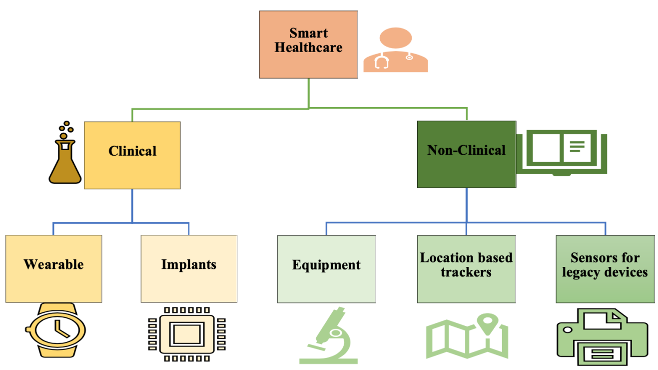


**Fig. 1. IoT based sensors in Health Care 4.0**

of Things (IoT) have recalibrated endless applications within the structure. The current generation, healthcare 4.0 improves clinical treatment such that medical practitioners can monitor personal health information shared through sensors to be more watchful and connected with the patients proactively. The smart IoT based devices available in the market have helped patient management by remotely monitoring health conditions and timely alerting the hospital about any irregularities using biomarkers on daily basis [3]. Smart healthcare, as shown in figure 1, works both on clinical and non-clinical data. Clinical trials for any disease require the subject to visit hospital and always be available physically for examination. With the help of IoT-based sensors, the health condition of the user can be tracked remotely using wearable IoT such as a wristwatch, or with the implantation sensor in the subject's body like a pacemaker. In the non-clinical collection of data, the bio-signals of the subject can be traced with the help of their smart devices such as mobile phones, the daily/ monthly activities like walking, running, and sitting to track the health of the user. Indeed, this health data fetched from IoT devices can allow caregivers to make informed decisions and therefore deliver better outcomes. The benefits of using IoT in healthcare include, but are not limited to:

- Higher patient engagement
- Better patient outcomes
- Decrease in errors
- Enhanced patient experience
- Timely intervention and diagnosis
- Improved accuracy
- Proactive treatments
- Better treatment outcomes

Mental healthcare also needs various biomarkers to detect the status of a person's mental health by evaluating the daily activities. An individual's behavior needs to be evaluated in different scenarios like his feeling while watching a movie, while driving, or while doing office work. Only then a psychologist can identify the actual mental condition of the person. Each psychological disorder has its own characteristic symptoms and some general warning signs to alert the need of professional help. An intelligent mental illness diagnostic can support clinicians with early detection. Most of the work conducted to detect the mental health status of a user involves datasets constructed with the help of medical questionnaires [4] [5]. These suffer due to lack of standardization of questionnaires, dishonest & unconscientious answers, unanswered questions and differences in understanding and interpretation. With healthcare 4.0, wearable IoT utilities can gather information, assess activity and other biomarkers, and even deliver interventions for various mental health conditions of an individual. Devices like wristwatch can be used to track the level of stress in an individuals [6]. It can collect the information about the daily activities of the subject such as walking, running, cycling along with bio-signals like the heartbeat, temperature, blood pressure and detect signs and symptoms of stress and direct the smartwatch wearer to therapeutic resources.

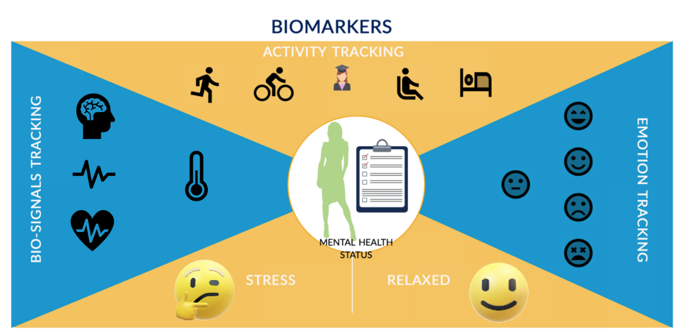


**Fig. 2. Mental Health Predictor Biomarkers**

In this work, we have proposed a model to detect the metal stress at an early stage by evaluating the different biomarkers indicative of mental health. As every individual act differently on encountering the same situation, stress is not dependent on a single attribute, and it works differently for different individuals. Therefore, to evaluate the stress level of an individual without prior medical history is a difficult task [7]. Various biomarkers can be used to track the mental health of an individual like sleep pattern, level of cortisol and adrenaline hormones, walking pattern, outdoor activities, size of eye pupil, heartbeat rate while performing physical activities and while in the resting period as shown in figure 2.

In this paper, we have proposed a hierarchal deep neural network that takes as input wearable stress and affects detection dataset (WESAD) that contains the bio-signals of 15 individuals collected from the wrist-wearable device (Empatica E4) and chest-worn device (RespiBAN) for a time-span of 2 hours. The different biomarkers that are taken into account to identify the stress in a user include Electrocardiogram (ECG) signals, Body temperature (TEMP), Blood Volume Pulse (BVP), Electrodermal Activity (EDA), Respiration, Three-Axis Acceleration (ACC) motion, Electromyogram (EMG). WESAD was recorded in the lab under a controlled environment by showing different visual stimuli and by giving assignments to evaluate the behavior of a user at that time. A total of 16000000 instances of data are available. The mental state of the subject is categorized into three classes, namely, baseline, stress, and amusement. Also, the bio-signals are a collected as a time series data. So, to evaluate the mental state of the individual, input is a frame of 1 second with the sliding window of 0.25 second, since the effect of an external stimuli can change over 5 seconds.

The hierarchal deep neural network consists of three levels. As WESAD contains data of each device-type bio-signal separately, so, to fetch the optimal values for every feature at each instance, at the first level, the sub-sub networks(SSN) for each bio-signal are used. Each bio-signal based sub-sub network (10 SSNs) is a 1-dimensional convolution neural network (1D-CNN) containing two convolution layers with batch normalization and max-pooling and one dense layer. These SSNs generate a high-level representation of the respective wrist and chest-based biomarkers which are input to the respective SNs at the second level. Again, both these sub-networks at the second level are 1-dimensional convolution neural network (1D-CNN) containing two convolution layers with batch normalization and max-pooling and one dense layer. This produces a combination the high-level representation features of each device type biomarker. Finally, at the third level, separately learned device type biomarkers are combined into one unified representation realizing a model-level fusion strategy. Thus, the shared representation is given to a convolution layer which generates the final feature vector and uses a denser layer to process the feature vector. The output layer generates a regression output with linear activation to finally detect the mental state into one of the three categories, baseline, stress, and amusement. The network is compared with the machine learning (Decision tree and Random forest classifier, LDA, KNN, AdaBoost), and deep learning models (CNN with late fusion), which is the state-of-the-art result for the WESAD dataset proposed by Schmidt et al. in 2018 and Lin et al. in 2019. The primary contributions of this research are:

- A deep neural network with hierarchical learning capabilities is proposed for mental stress state detection.
- Both wrist-based and chest-based sensor bio-signals are used to generate high-level representation of features with generalization capabilities.
- Model-level fusion strategy is proposed to elucidate the correlation in data as different sub-networks are used to operate over features which are learned separately for each input type and then combined into one unified representation.

### *1.1. Related Work*

Mental health is as vital as physical health. Most of the organizations try to arrange motivational sessions or activity sessions to ensure their employees get a 'mental vacation' and feel relaxed. This is important as a stressed or depressed person will always find it challenging to focus and so will always take extra time to pursue the same work in comparison to a mentally healthy and relaxed person. As the early stages of mental illness have invisible symptoms, it is not always possible to notice the change until the symptoms are persistent, increase in frequency and severity and interfere with life activities and roles. By this time, it becomes too late and in worst cases, untreated mental illnesses can lead to loss of life an average of 25 years early [5]. Thus, early identification and intervention are necessary to recover and reclaim lives. Various artificial intelligence based techniques have been reported to supplement clinical practice in various mental healthcare studies. To analyze the behavior of a person under stress, researchers have proposed machine learning-based techniques; however, the availability of a WESAD dataset has always been an issue. In 2018, Schmidt [8] created and applied five different machine learning algorithms namely, Random Tree Classifier (RT), Decision Tree (DT), AdaBoost, K Nearest Neighbor (KNN), and Linear Discriminant Analysis (LDA) and given the state of the art results by providing the accuracy of 80%. In 2019, Lin et al. proposed a deep fusion network on the WESAD to optimize the accuracy of the prediction of stress. They used the late fusion method in deep neural networks and divided the model into four subnetworks, one tuned on the chest sensor dataset, and rest three tuned on the wrist sensor dataset, first one on EDA and Temperature, second on BVP and last on ACC. They attained the highest accuracy of 85% and F1 score as 0.86, which was a significant improvement on the result provided by Schmidt et al. [9]. The latest work on stress detection using WESAD is proposed by [10]. They used three classifiers, namely logistic regression, decision tree, and random forest, and rather than evaluating the result into three categories, they added one more output category as meditation. Also, rather than applying each classifier on the complete dataset, they applied it on individual subjects, resulting in an accuracy of 88% to 99% for the individual subject.

WESAD is considered as the most recent benchmark dataset to analyze mental health of a person since it contains the maximum number of biomarkers on a single subject to determine the affective state of the subject. Previously, many researchers have created datasets to evaluate the stress level. Picard et al. [11] built a dataset containing physiological data of a single person depicting eight different emotions for 20 days. As different individuals can represent different behaviors for the same emotional feeling; therefore, the dataset collected from a single source cannot predict accurately for all the users. Healey et al. [12] also created a dataset for evaluating stress using ECG, Electrodermal Activity and Respiration, and Electromyogram data. However, this dataset was only used to evaluate the stress of a driver, so it did not apply to all the subjects performing different actions. In 2012, DEAP was published by Koestra et al.; which contained the facial videos and EEG signals of

the users to analyze the emotions using peripheral signals [13]. Although it used multiple subjects, the features used were limited, so DEAP was also unable to predict the emotion of all the users accurately. [14] used mobile phones to construct data for analyzing the level of stress in a user. They used biomarkers like physical activity level, social interaction, and social activity along with the location of the user. A transfer learning model was used to predict the stress level and an accuracy of 76% was reported. In 2017, Gjoreski et al. [6] used only bio-signals from a wristwatch to evaluate the level of stress from 5 subjects evaluated over a span of 55 days using machine learning.
In the next section, we discuss the dataset used along with the model proposed. It is followed by a discussion on results and finally the conclusion is presented in the last section.

## 2. Material and Methods

Stress is a mental state of a person that can be triggered by external or internal stimuli, and these stimuli vary for different individuals. Consequently, to train a model, we need to have a variation of users with different scenarios. WESAD comprises data from 15 graduate students most prone to stress, 12 being male and three being female, are analyzed using both wrist-based and chest-based tests under diverse conditions over 2 hours. WESAD acts as a benchmark dataset to build the model. The prediction of the level of stress in a user depends on several biomarkers like dry mouth, dilation of pupils, decreased digestion, but measuring every aspect from an individual is not possible, to monitor the pupil dilation, we need access to video of the user continuously. However, it invades the privacy of the user. Thereby, data collection is done through sensors in a way that does not invade the privacy of the users.

### *2.1. Dataset*

To identify the mental stress using behaviour biomarkers, we have used a sensor collected multimodal dataset which features psychological and motion data from both a wrist-worn device Empatica E4, and chest-worn device RespiBAN over 2 hours in a lab under controlled environment named WESAD. It is a publicly available data set for variable stress and affects detection collected by Schmidt et al. in a lab study on 15 subjects 12 being male and 3 been female. Different biomarkers that were used hey to monitor the stress level of a person are Blood Volume Pulse (BVP), Electrocardiogram (ECG), Electrodermal Activity (EDA), Electromyogram (EMG), Body Temperature (TEMP), Respiration (RESP) and Three-Axis Acceleration (ACC) motion. The data collected by RespiBAN was at 700Hz, whereas the data collected by the wrist device was at low resolution. All these biomarkers contribute to identifying the mental state of a person, whether he is amused, stressed, or is in normal state i.e., baseline. Each subject has 12 features, and the results were self-reported by the users [8]. The dataset contains a total of 63000000 instances [8].

**Table 1. Biomarkers of the Proposed Model**

| Biomarker | Device | Feature/ Significance |
|---|---|---|
| Blood Volume Pulse | RespiBAN, Empatica E4 | BVP is the amount of blood in blood tissue during a certain time period. BVP also provides pulse rate and blood flow volume, as it is obtained by photoplethysmography. |
| Electrocardiogram | RespiBAN | ECG provides the frequency of cardiac cycles. It is sensed using photodetectors, so is not able to be detected by wrist-wearable device. |
| Electrodermal Activity | RespiBAN, Empatica E4 | EDA gives the flow of electricity through the skin. The changes arise in skin when brain sends the signal due to different emotion activation. Skin conduction increases when a person is under stress. |
| Electromyogram | RespiBAN | EMG is used to detect musculoskeletal movements. These signals can detect face and hand gestures. |
| Body Temperature | RespiBAN, Empatica E4 | Skin temperature of the subject is measured using thermistor sensor. Body temperature is negatively correlated with stress. |
| Respiration | RespiBAN | RESP gives the person inhalation and exhalation rate. The slowed respiration rate shows the level of stress in a user. |
| Three Axis Acceleration | RespiBAN, Empatica E4 | ACC gives an indication of different activities like lying, sitting, standing, walking, running and cycling by recording the human movement in all the three dimensions. Fast hand movement over short time depicts sign of mental stress. |

### *2.2. Convolution Neural Network (CNN)*

Convolution Neural Network (CNN) is the most commonly used deep neural network proposed for image processing but now validated for all types of data. Typically, in a CNN, the input is passed through convolution layers such that the output of the primary layer becomes the input for the subsequent layer. Non-linearity is added post every convolution operation using an activation function such as ReLU, to create a rectified feature map. Each non-linear layer is followed by a pooling layer which performs a down sampling operation. Pooling operation helps to progressively reduce the size of the input representation and control overfitting too [15]. We can either use max, average or sum pooling. A fully connected layer also known as the dense layer is then attached to this series of convolution, non-linear and pooling layers which outputs the information from the convolutional networks. Structure of CNN model is shown in figure 3.
In our proposed hierarchical network we have three levels. At bottom level, we use 1-D CNN for creating 10 different sub-sub networks(SSN) to generate the optimized feature value for every bio-signal of RespiBAN and Empatica E4. In the second layer, the output of SSNs are passed to two different sub-networks(SN) depending upon the feature is obtained by RespiBAN or Empatica E4, for producing a combination of the high-level representation features of each device type biomarker. Both SNs are also a 1D-CNN containing two convolution layers with batch normalization and max-pooling and one

dense layer. And the top level is the classification level, where the result of SNs, separately learned device type biomarkers are combined into one unified representation realizing a model-level fusion strategy. As the dataset has approximately more than half-million values for a single attribute over a single subject, the fusion strategy is used. Typically, fusion strategies can be categorized into early, model-level and late fusion. The early fusion strategy involves concatenation of input features whereas the model-level fusion involves concatenation of high-level feature representations from different sub-networks and the late multi-lingual fusion involves fusion of predictions from different sub-networks (figure 4).

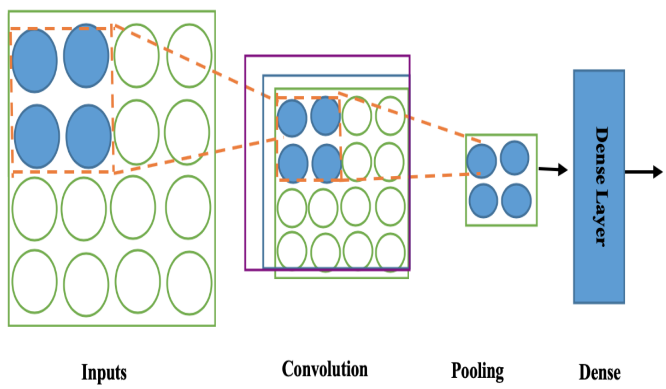


**Fig. 3. CNN Structure**

The model-level fusion strategy helps to complete the essence of hierarchical learning proposed in this work. Unlike early fusion, this strategy helps to circumvent the curse of dimensionality and synchronization between different features and at the same time does not isolate interactions among different features as in late fusion. Finally, the fused representation is given to a convolution layer, which generates the final feature vector and uses a denser layer to process the feature vector. The output layer, i.e. SoftMax layer generates a regression output with linear activation to finally detect the mental state into one of the three categories, baseline, stress, and amusement.

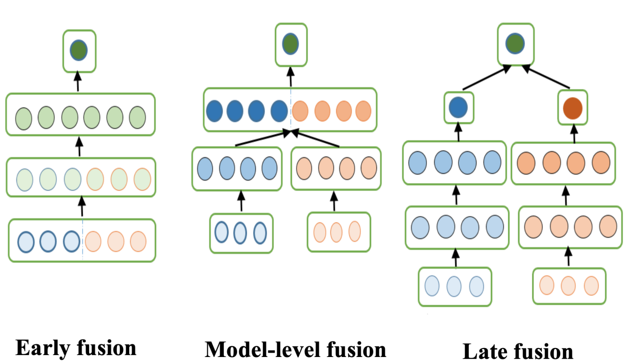


**Fig. 4. Fusion Strategies**

## 3. Model

### *3.1. Data Pre-Processing*

Since the different subjects have a different response to the different type of tests, they have variation in their signal values, to use this dataset we apply the min-max normalization technique so that all the subjects have test results in the same scale range. Since the data collection is done over 2 hours, in which a subject is presented with different stimuli, it results in different behavior at different times. To evaluate a time series data, the continuous data is broken into short instances of one second. The sliding window of 0.25 seconds is used to better predict the mental state, as the effect of stimuli can happen over even a short interval of 5 seconds. Over breaking the continuous time frame data, if the change of stimuli started in previous one second interval, but the main change comes in the next one second interval then to compare the change, the value of previous instance is required, that can be used with the sliding window of 0.25 second, which is to use the prior time to predict the next time step.
The target class is also given the numerical value, 1 for baseline, 2 for stress, and 3 for amusement; Baseline is a condition that a subject exists in the first 20 minutes of the experiment setup, where no external stimuli has been used. For all those scenarios, where target class is undefined is given the value of 0.

### *3.2. Architecture*

After the data has been fetched in the scalable format, data is still astronomical and exists as individual units, so we apply the 1D-CNN over each feature signal by constructing a separate sub-sub network. Once all the sub-sub network gets trained. The input is passed in as subject manner, i.e., all the features of a subject at one sliding window or at a particular time is passed through the sub networks, and the output of the sub models is then passed into the primary model that classifies the subject at a particular instance into one of the three classes defined during the training phase. The proposed hierarchical network is shown in figure 5.
Each subject's data contains signals from two different sensor devices, first RespiBAN, for monitoring chest signals ECG, TEMP, RESP, EDA, EMG, and ACC each recorded with 700Hz signal, the sub-sub-networks created for each of these feature is 1-D CNN with input layer 700*1 except for ACC which is provided with input layer of 700*3. The sub-sub-networks for features recorded with Empatica E4, have different sized input layers, for ACC 32* 3, BVP has 64*2, TEMP and EDA has input size of 4*1 input layer. 65% of data is used to train the model, i.e., ten subjects are used to train the model, and rest 35% is used to test the data, and 10% of the data is used to validate the model. 53% of the total instances belong to the baseline class, 30% belongs to stress class, and 17% belong to the amusement class. The structure of different layers used in CNN sub-networks implemented in the proposed hierarchal network is shown in table 2.

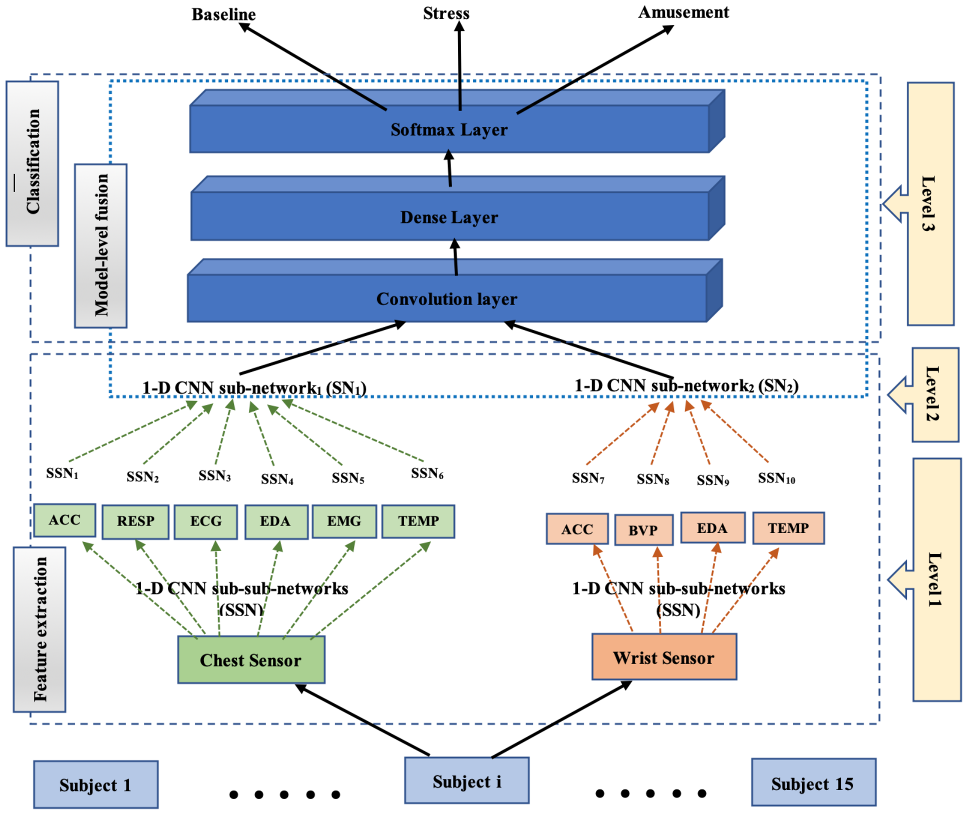


**Fig. 5. Architecture of Proposed Deep Hierearichal Model**

**Table 2. Layers of the proposed Hierarchical deep neural network**

| | CNN-SSN1 to CNN-SSN10 | CNN SN1 and CNN SN2 |
|---|---|---|
| High-level Feature Extraction | Input | Input |
| | Convolution | Convolution |
| | Pooling | Pooling |
| | Convolution | Convolution |
| | Pooling | Pooling |
| | Dense | Dense |
| Classification based Model-level fusion | Convolution | |
| | Dense | |
| | SoftMax | |

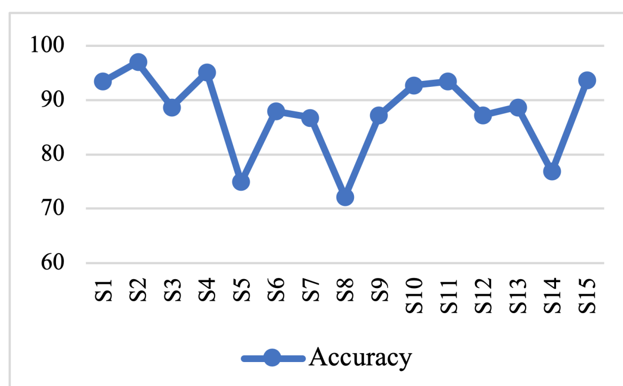


**Fig. 6. Accuracy of Subjects**

## 4. Results & Discussion

The proposed deep hierarchal neural network is trained with 65% of processed WESAD data once the continuous time series data is split over one second interval over the sliding window of 0.25 second. Model is tested over 35% of total data, i.e. 5 subjects are used for testing the models. Model is evaluated in terms of accuracy and f1-Score, where,

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{F-1 Score} = 2 * \frac{Precision*Recall}{Precision+Recall}$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{and Recall} = \frac{TP}{TP+FN}$$

Precision, Recall, Accuracy and F1-score of each individual subject is hsown in table3. The average accuracy achieved by the proposed model is better than the state-of-the-art results. The F-score of the model is 0.8325, and the average accuracy is 87.7% which is better than the state-of-the-art results provided by [8],[9]. The comparison of the deep hierarchal model with results of other models is shown in table 4.

The accuracy of the model depends upon the subject's data, as different individuals have different level of stress for same scenario and the expression of stress is also different for each individual. The results of the proposed model for each subject is shown in table 4. It is observed that the accuracy of the model varies from 72% to 96%. The accuracy curve of model is shown

**Table 3. Accuracy of different classifiers over WSEAD**

| F-1 Score | Precision | Recall | Accuracy | Subjects |
|---|---|---|---|---|
| 0.9467 | 0.901 | 0.9973 | 93.39 | **S1** |
| 0.998 | 0.9474 | 0.9861 | 96.98 | **S2** |
| 0.612 | 0.826 | 0.9193 | 88.70 | **S3** |
| 0.9292 | 0.9079 | 0.9693 | 95.07 | **S4** |
| 0.96855 | 0.5518 | 0.9051 | 74.9 | **S5** |
| 0.8548 | 0.7709 | 0.9593 | 87.92 | **S6** |
| 0.8601 | 0.8353 | 0.8877 | 86.79 | **S7** |
| 0.6836 | 0.8761 | 0.5605 | 72.19 | **S8** |
| 0.798 | 0.8145 | 0.8756 | 87.24 | **S9** |
| 0.9348 | 0.9921 | 0.8831 | 92.73 | **S10** |
| 0.834 | 0.873 | 0.913 | 93.46 | **S11** |
| 0.8577 | 0.8222 | 0.8974 | 87.24 | **S12** |
| 0.93 | 0.92 | 0.961 | 88.72 | **S13** |
| 0.649 | 0.789 | 0.681 | 76.84 | **S14** |
| 0.917 | 0.95 | 0.925 | 93.56 | **S15** |

**Table 4. Accuracy of different classifiers over WSEAD**

| Classifier | Accuracy | F-score |
|---|---|---|
| **Decision Tree** [8] | 0.64 | 0.58 |
| **Random Forest** [8] | 0.75 | 0.64 |
| **KNN** [8] | 0.56 | 0.48 |
| **LDA** [8] | 0.75 | 0.71 |
| **AdaBoost** [8] | 0.79 | 0.69 |
| **CNN** [9] | 0.85 | 0.86 |
| **Proposed Hierarchical Model** | 0.877 | 0.83 |

in figure 6. The sample confusion matrix for subject 5 and 10 is shown in figure 7.

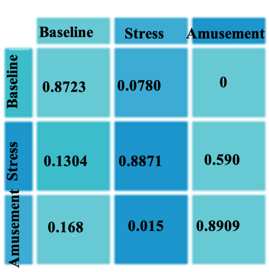


**Fig. 7. Subject 5**


**Fig. 8. Subject 15**

## 5. Conclusion

Bio-signals are the biomarkers depicting these physiological changes as stress response symptoms during chronically activated situations. Only trained medical practitioners can measure such indicators, which can be tedious and time-consuming, thus delaying early identification and timely intervention. Persistent and chronic stress can lead to long-term health damage. It is imperative to design and develop an intelligent mental illness diagnostic that can support clinicians pro-actively. With the availability of smart sensors, the health condition of a person both physically and mentally can be tracked easily through [1] IoT based wearable devices. This research proffered a deep hierarchal neural network model which, on receiving different sensor-based signals, categorizes the individual mental state to be either in stress or not or in amusement. A multivariate time series data, wearable stress and affects detection dataset (WESAD), consisting of both wrist-based and chest-based sensor bio-signals, was trained using a hierarchy of networks. The model used 13 CNN networks at different levels of the hierarchy that provided an average accuracy of 87.7%, which is more efficient than the state-of-the-art results.

## 6. Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

1. J. L. P. S. M. Bienertova-Vasku, 'Eustress and Distress: Neither Good nor Bad, but rather the same?', 2020.
2. S. Kumar, 'Biochemistry of the Brain', in Examining Biological Foundations of Human Behavior , 2020.
3. S. D. a. R. B. Rustem Dautov, 'Hierarchical data fusion for Smart Healthcare', Journal of Big Data, pp. 1-23, 2019.
4. L. Z. Q. S. A. L. L. &. Z. H. Tong, 'Inverse boosting pruning trees for depression detection on twitter', 2019.
5. A. E. A. M. &. V. C. ,. Esposito, 'Needs and challenges in human computer interaction for processing social emotional information ', Pattern Recognition Letters, pp. 41-51, 2015.
6. M. L. M. G. M. ,. G. H. Gjoreski, 'Monitoring stress with a wrist device using context', Journal of Biomedical infermatics, pp. 159-170, 2017.
7. S. F. M. C. A. C. A. &. F. G. Barra, 'PhysioUnicaDB: A dataset of EEG and ECG simultaneously acquired ', Pattern Recognition Letters, pp. 119-122, 2019.
8. P. R. A. D. M. C. a. L. K. Schmidt, 'Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection', in International Conference on Multi- modal Interaction (ICMI '18), Boulder, CO,USA, 2018.
9. J. P. S. L. C. S. O. S. Lin, 'An Explainable Deep Fusion Network for Affect Recognition Using Physiological Signals', in 28th ACM International Conference on Information and Knowledge Management (CIKM 19), Beijing, China, 2019.
10. F. I. ,. W. S. Indikawati, 'Stress Detection from Multimodal Wearable Sensor Data', in 2nd International Conference on Engineering and Applied Sciences (2nd InCEAS), 2020.
11. R. V. E. a. H. J. Picard, 'Toward machine emotional intelligence: Analysis of affective physiological state', IEEE Transactions on Pattern Analysis and Machine Intelligence , pp. 1175-1191, 2001.
12. J. ,. P. R. Healey, 'Detecting stress during real-world driving tasks using physiological sensors', EEE Transactions on Intelligent Transportation Systems, pp. 156-166, 2005.
13. S. M. C. ,. P. I. Koelstra, 'Deap: A database for emotion analysis; using physiological signals', IEEE Transactions on Affective Computing, pp. 18-31, 2012.
14. P. H.-L. L. E. S. ,. O. V. E. F. M. ,. O. M. Alban Maxhuni, 'Stress modelling and prediction in presence of scarce data', Journal of Biomedical Informatics, pp. 344-356, 2016.
15. P. R. D. D. P. L. D. B. G. S. D. N. M. R. S. &. F. M. Melillo, 'Wearable improved vision system for color vision deficiency correction ', IEEE journal of translational engineering in health and medicine, pp. 1-7, 2017.
16. J. K. a. E. André, 'Emotion recognition based on physiological changes in music listening', IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 2067-2083, 2008.
17. A. K. A. a. S. M. Zenonos, 'HealthyOffice: Mood recog- nition at work using smartphones and wearable sensors', in PerCom Workshops., 2106.
18. Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. , 'Hierarchical Attention Networks for Document Classification', in 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, 2016.
19. Lo, S. C. B., Chan, H. P., Lin, J. S., Li, H., Freedman, M. T., & Mun, S. K. 'Artificial convolution neural network for medical image pattern recognition', Neural networks, 1201-1214,1995.