This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

Non-invasive Hydration Level Estimation in Human Body using Galvanic Skin Response

Ali Rizwan, *Student Member, IEEE,* Najah Abu Ali, *Senior Member, IEEE,* Ahmed Zoha, *Member, IEEE,* Metin Ozturk, *Student Member, IEEE,* Akram Alomaniy, *Senior Member, IEEE,* Muhammad Ali Imran, *Senior Member, IEEE,* and Qammer H. Abbasi, *Senior Member, IEEE,*

***Abstract*—Dehydration and overhydration, both have mild to severe medical implications on human health. Tracking Hydration Level (HL) is, therefore, very important particularly in patients, kids, elderly, and athletes. The limited solutions available for the estimation of HL are commonly inefficient, invasive, or require clinical trials. Need for a non-invasive auto-detection solution is imminent to track HL on a regular basis. To the best of authors' knowledge, it is for the first time a Machine Learning (ML) based auto-estimation solution is proposed that uses Galvanic Skin Response (GSR) as a proxy of HL in the human body. Various body postures, such as sitting and standing, and distinct hydration states, hydrated vs dehydrated, are considered during the data collection and analysis phases. Six different ML algorithms are trained using real GSR data, and their efficacy is compared for different parameters (i.e., window size, feature combinations etc). It is reported that a simple algorithm like K-NN outperforms other algorithms with accuracy upto 87.78% for the correct estimation of the HL.**

***Index Terms*—Skin Conductance Level (SCL), GSR, Electrodermal Activity (EDA), Hydration Level, Machine Learning (ML), Bio-Sensors Data.**

I. Introduction

Dehydration and overhydration both are associated with morbidity and mortality. Dehydration may lead to gastrointestinal, urological, metabolic, circulatory, neurological disorders or fatigue and overhydration may cause edema, hyponatremia etc. [1]. In a developed country like the USA alone, more than 1.5 million children are diagnosed with acute diarrhea, a gastroenteric disease, and approximately 300 of them die annually. In developing countries, it causes around 2 million deaths annually and it is a very common cause of death in children under the age of five years. According to World Heath Organization (WHO) approximately 4 billion cases of diarrhea are reported worldwide annually. Dehydration is also fatal for elders, 30-day mortality with a principal diagnosis of dehydration is 17% in elderly patients which approaches to 50% at the one-year mortality rate [1]. Maintaining appropriate Hydration Level (HL) is not only important to avoid diseases but it is also equally important for healthy people to perform the activities of daily life. Water is the major component of blood which transports nutrients from food and oxygen from lungs to the cells of the body.

To maintain an appropriate HL, there is a need for an HL monitoring system. Literature review shows some parameters for the measurement of HL in the human body but a gold standard for the quantitative comparison is still missing. The commonly used parameters include Plasma Osmolality (PO), Total Body Water (TBW), Bioelectrical Impedance (BIA), Urine Content and Salivary Flow Rate. However, the challenge is that the most of methods available are invasive in nature, such as the use of isotopic dilution or they involve clinical trial like common techniques of BIA [2]. Basic purpose of BIA is to assess the body composition for the estimation of Fat Mass (FM) and Fat-Free Mass (FFM). Additionally, it can also provide information about TBW. BIA is a noninvasive method that comprises the measurement of resistance to a weak electric current passed through the body. That resistance level is then used to estimate FM, FFM and TBW [3]. The drawbacks of using BIA are, it is an on-demand complex method that can not be used for continuous monitoring, and it is an indirect measure of HL.

Nevertheless, from the solutions like BIA, it is evident that there exists a correlation between HL and electric resistance of the body. The electric resistance of the human body varies from few ohms to thousands of ohms depending on the water content in the body. More than 99% of the resistance against the flow of electric current through human body is faced at skin level [4]. Measurement and analysis of Galvanic Skin Response (GSR), a measure of resistance or conductance of the skin, can help to develop a noninvasive wearable solution for HL monitoring. Automated monitoring and estimation of HL is crucial for a plethora applications in healthcare. It is, therefore, attracting more research in different HL monitoring approaches such as HL monitoring via body temperature [5], tracking the activity and water consumption [6], and using skin impedance [7].

Previously, in a preliminary study, HL is detected using skin conductance data of single participant and a feature space comprising only three features. In that study HL is detected in sitting, standing and posture independent scenarios with an accuracy of 73.91%, 81.82%, and 60.71%, respectively [8]. As extension of that work, here, in this study, a more systematic and detailed analysis is performed. Now, more generalized model is developed exploiting data of multiple participants and a larger feature space. This model is not only more generalized

A. Rizwan, A. Zoha, M. Ozturk, M. A. Imran, and Q. H. Abbasi are with the James Watt School of Engineering, University of Glasgow, Glasgow, UK e-mail: (a.rizwan.1@research.gla.ac.uk; ahmed.zoha@glasgow.ac.uk; muhammad.imran@glasgow.ac.uk; qammer.abbasi@glasgow.ac.uk).

N. A. Ali is with the College of Information Technology, United Arab Emirates University,Al Ain, UAE e-mail: (najah@uaeu.ac.ae).

A. Alomaniy is with School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK e-mail: (a.alomainy@qmul.ac.uk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

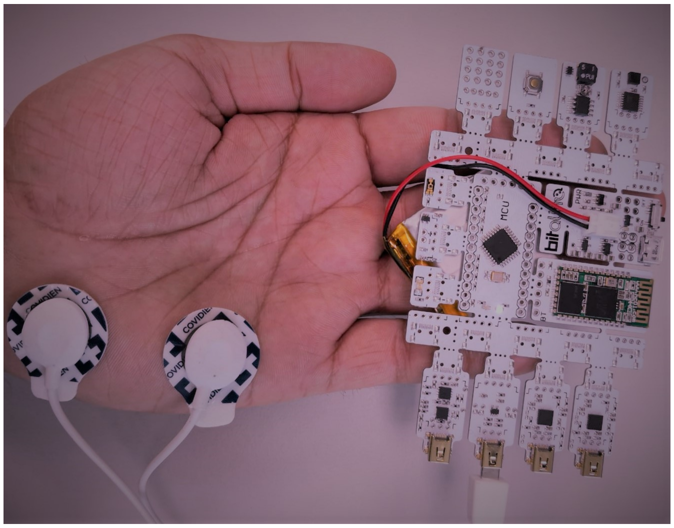


Fig. 1. GSR data collection using BITalino EDA sensor by placing two electrodes on the hypothenar side of the palm of the non dominant hand.

but it also yields much better accuracy. To the best of authors' knowledge, for the first time in the literature this paper is presenting GSR also known as Skin Conductance Level (SCL), or Skin Resistance Level (SRL), as a proxy for the auto-estimation of the HL in human body. For the measurement of GSR, Electrodermal Activity (EDA) sensor is used which basically measures the electrical resistance of the skin. EDA is mainly used for the study of sympathetic behaviour in humans. GSR is very sensitive to the electric signals produced by the nervous system as the result of any internal or external stimuli [9]. Here, it is proposed that GSR can be used as an indicator of HL in the body. It is so because the GSR reflects the variations in the skin resistance and conductance which are also correlated with water content in skin. Skin conductance increases with the water content in the skin. On the other hand, skin water content is also correlated to the overall HL of the body [10]. One challenge in the usage of GSR, measured by EDA sensor, is that it is very sensitive to internal and external stimuli for the nervous system. Stimuli such as happiness, fear, temperature, humidity etc., can lead to rapid variations in GSR measurements. Whereas variations in HL are not an impulsive behaviour. This issue can be resolved by measuring GSR for longer intervals of time and using features for tonic activities rather than the phasic activities.

Major contribution of this study is the development of a Machine Learning (ML) model for the auto-estimation of HL in the human body using GSR data. To achieve that goal such feature set, window size and algorithm are identified that together give the best performance for the HL estimation. To find the best model for HL detection a comparative study is performed on the GSR data collected in different body posture based scenarios. Study involves analysis of the performance of multiple state-of-the-art ML algorithms applied over numerous feature sets extracted from the data segments of different sizes (i.e data collected for different length of time).

The rest of the paper is organized as follows: Section II describes the data collected, while Section III explains the methodology from data collection to model optimization. The results are presented in Section IV, and analysis and discussion on the findings are covered in Section V. Lastly, Sections VI concludes the paper.

TABLE I
COUNT OF DATA SAMPLES AFTER FEATURE EXTRACTION FOR DIFFERENT WINDOW SIZES OF TIME, BODY POSTURES AND HYDRATION STATES

Posture	**State**	**Number of Samples for Window Size**			
		30 sec	**45 sec**	**60 sec**	**75 sec**
Stand	**Hydrated**	290	205	150	125
	Dehydrated	290	205	150	125
Sit	**Hydrated**	300	200	150	120
	Dehydrated	300	200	150	120
Independent	**Hydrated**	640	425	325	260
	Dehydrated	640	425	325	260

II. DATA DESCRIPTION

Data used in this research is collected from five individuals. The dataset comprises the data from four males and one female of different ethnicity, all in the bracket of 25 to 30 years of age. None of the participants has any known conditions of edema or hypo-hydration. Data is collected in two states labelled as hydrated and dehydrated. The participants are considered dehydrated when they have not had any intake of fluid or food for the last 10 hours at least. For the hydrated state data is collected within an hour of the intake of plenty of water and when the participants have been drinking water frequently earlier.

Moreover, both hydrated and dehydrated state data are collected in two scenarios, based on physical postures, sitting and standing. Another data set is created by combining the data of both postures and it is labelled as posture independent scenario data. In total, it is data of around 10 hours, 2 hours of data for each individual. Out of those 2 hours, the data of 1 hour is measured in hydrated state and the rest in dehydrated state. Out of that 1 hour, data of half-hour is collected in sitting posture and that of other half-hour is collected in standing posture. Data is recorded in samples of 5 to 15 minutes in all scenarios to avoid any issues due to sweating on the palm.

The data is collected at a resolution of 16 bit and a sampling rate of 1 MHz (i.e., 1000 samples every second), the highest available precision options on the BITalino kit [11]. The kit computes GSR as

$$GSR = \frac{1}{R}, \tag{1}$$

where R is the skin resistance in MΩ:

$$R = 1 - \frac{C}{2^n}, \tag{2}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

where C is the value sampled from the channel of BITalino kit at resolution n. GSR, measured in μS, is defined as a skin conductance level and/or the reciprocal of the skin resistance.

It is worth mentioning that the raw GSR signal data is not used directly in ML algorithms. Instead, statistical features are extracted from the raw data for different length of time intervals called windows of time. Count of samples in data sets created after the feature extraction and used for training and testing are presented in Table I.

III. Proposed Methodology

A salient feature of the model development approach in this study is the implementation of parallel processing in two stages as presented in Algorithm 1. In the first stage, each data set (D) is split into small segments (smaller data sets) based on four different windows of time (W) and then, features (F) are extracted from each segment. Once all features are extracted for every data set and window size then new data sets (D_{f}) are created comprising the feature vectors. A summary of feature data sets is presented in Table I. Next in the second phase, six different ML algorithms are trained and evaluated for each feature data set separately, again in parallel. Apart from that, the important steps of the methodology followed in this study are summarized in Figure 2 and illustrated briefly as as follows:

A. Data Collection

Formal approval is obtained from the competent authority prior to the data collection and data is collected following the guidelines provided on the EDA data collection by the ad-hoc Committee of the Society for Psychophysiological Research on Electrodermal Measures [12]. A BITalino EDA sensor kit, as shown in Figure 1, is used for the GSR data collection.

Algorithm 1: Model Development based on Parallel Processing

1 **Data Sets** (D) $\in$ {sitting (D_{sit}), standing (D_{sd}), independent (D_{ind})}
2 **Features** (F) $\in$ {minimum, maximum, variance, entropy, standard deviation, percentile, median, mode, kurtosis}
3 **Window Sizes** (W) $\in \{30,\ 45,\ 60,\ 75\}$ seconds
4 **Algorithms** (A) $\in$ $\{$ K-NN, LR, DT, SVC, NB, LDA$\}$
5 **do in parallel**
6   Extract F from each data set in D for each window size in W
7   **return** *the data extracted in Step 6*
8 Create holistic feature data sets (D_{f});
9 **do in parallel**
10   Train a model for each algorithm in A with cross-validation on D_{f}
11   Test the models developed in Step 10
12   **return** *the best models for each data set in D*

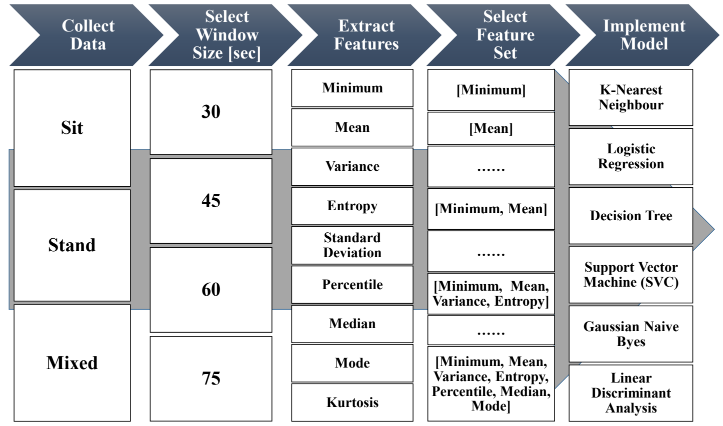


Fig. 2. Main steps of proposed methodology for the model development.

The readings of BITalino EDA sensor kit basically show the skin conductance at the time of data collection. The data is collected for each participant in hydrated and dehydrated state in two body postures sitting and standing. Overall three data sets are analyzed based on the body posture scenario, named as standing posture scenario, sitting posture scenario and posture independent scenario.

B. Window Size Selection

Initially the data is collected for longer intervals like 5 to 30 minutes. A window operation is applied to split data into smaller segments of shorter windows of time (i.e. intervals of time). Features are then extracted from those segments. Given that different window sizes exhibit different underlying data patterns, an important task is to identify an optimal window size. Four window sizes $W \in \{30, 45, 60, 75\}$ are used in this experiment and each window size presents the time interval in seconds. Once the optimal window size is determined, the window operation can also help to estimate HL for the data being collected at real-time. In this study, an overlapping window is not used as tonic characteristics of the GSR data are more relevant rather than the phasic ones [12]. It is so because HL is not expected to be an impulsive behaviour.

C. Feature Extraction

A feature space, F, of following nine statistical features is used: $F \in$ {Minimum, Mean, Variance, Entropy, Standard Deviation, Percentile, Median, Mode, Kurtosis}. The values of each feature, from the feature space, is computed for every window size. For example, when a window size 30 seconds is selected, data is split into non-overlapping segments of 30 seconds and statistical features are computed for each segment of 30 seconds. At the end of feature extraction, new data sets are created comprising nine vectors, each of them contains values for each statistical feature in F. The number of samples or entries in new feature data sets are presented in Table I.

D. Feature Selection

After feature extraction, another important task is to identify the combination of features that generate the best accuracy for the estimation of HL. An algorithms based parallel heuristic approach [13] is applied in which all possible combinations of features are evaluated for each algorithm. In total $(2^f - 1)$ combinations of features are evaluated for each posture based scenario. Even though it is a time consuming heuristic process which is also computationally expensive, the proposed parallel processing approach, presented in Algorithm 1, makes it feasible.

E. Implementation of ML Models

Six ML algorithms trained and evaluated in this study are Logistic Regression (LR), Support Vector Machine based Classifier (SVC), Decision Tree (DT), K-Nearest Neighbour (K-NN), Linear Discriminant Analysis (LDA), and Gaussian Naive Byes Classifier (NB). All the algorithms are separately evaluated for all the possible combinations of features for each window size and posture. In total $[a \times s \times w \times (2^f - 1)]$ number of models are evaluated to find the best model, where a, s, w, and f are the count of algorithms, data sets of postures based scenarios, window sizes, and data sets of features respectively. For every model, 70% of the feature set data is used for training and three-fold cross-validation, whereas 30% of the data is kept separate to test each model at the end.

Another critical task performed is optimization. Different algorithm-specific hyperparameters, explained in [14] and listed in Table II, are evaluated to fine-tune the model for the optimal results. For the rest of the less significant parameters, default values provided in [14] are used. Three fold cross validation is used to find the values of tuning parameters for the best performing model and to ensure the consistency of the trained model (i.e. the model is neither over-fitted nor under-fitted). Then the best performing algorithm is selected with the corresponding parameter values.

F. Model Evaluation

It is a classical binary classification problem as there are two states: hydrated and dehydrated. The performance of ML algorithms applied is assessed using some most popular metrics listed below [15]:

$$Precision = \left(\frac{t_{\mathrm{p}}}{t_{\mathrm{p}} + f_{\mathrm{p}}}\right), \tag{3a}$$

$$Sensitivity = \left(\frac{t_{\mathrm{p}}}{t_{\mathrm{p}} + f_{\mathrm{n}}}\right), \tag{3b}$$

$$Specificity = \left(\frac{t_{\mathrm{n}}}{t_{\mathrm{n}} + f_{\mathrm{p}}}\right), \tag{3c}$$

$$CCR = \left(\frac{t_{\mathrm{p}} + t_{\mathrm{n}}}{P + N}\right), \tag{3d}$$

where t_{p} is truly positive, the number of dehydrated instances detected correctly, and f_{p}, the false positive, represents the hydrated instances detected as dehydrated instances. f_{n}, false negative, is the number of dehydrated instances which have been wrongly detected as hydrated instances. t_{n}, truly negative, is hydrated instances detected correctly. P and N are the actual numbers of dehydrated and hydrated instances.

CCR, the correct classification rate also known as overall accuracy (%) is initially used to assess the performance of all the algorithms in all possible scenarios (i.e., for all window sizes, postures and combination of features). When the best performing models are shortlisted based on CCR score in different posture based scenarios then metrics like precision (recall), sensitivity (aka True Positive Rate (TPR)), Specificity (aka True Negative Rate (TNR)) are calculated to asses the detailed performance of those models.

IV. Results

As it can be seen from Figure 3, even a classical ML algorithm can identify HL from GSR data with accuracy upto 87.78%. Overall, K-NN outperforms the other algorithms for the posture specific and posture independent scenarios as shown in Fig. 3. For the feature combinations A, B, and C

TABLE II
Algorithms specific important hyperparameters and their values used in model optimization

Algorithm	Hyperparameter
KNN	Metrics: [Minkowski, Euclidean, Manhattan]
	Weights: [Uniform, Distance]
	K: x, 1 ≤ x ≤ n-2, where n= No. of Samples
LRA	Penalty = [L1,L2]
	C: [0.001,0.01,0.1,1,10,100]
DT	Criterion: [Gini,Entropy]
	Max depth:x, 3 ≤ x ≤ 25
SVC	Kernel: [Linear, Rbf,Poly]
	C: [0.1,1,10,100,1000]

TABLE III
Features Combinations with the best Performance for Different Window Sizes and Postures

ID	Combination	Window Size (sec)	Posture	Accuracy [%]
A	Mean, Variance, Entropy, Percentile Standard Deviation,	60	Sit	87.78
B	Mean, Variance, Entropy, Percentile, Standard Deviation, Median	60	Sit	87.78
C	Mean, Entropy, Standard Deviation, Percentile	60	Sit	87.78
D	Minimum, Mean, Entropy, Mode, Standard Deviation, Kurtosis	60	Stand	83.33
E	Minimum, Mean, Variance, Entropy	30	Independent	76.82

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

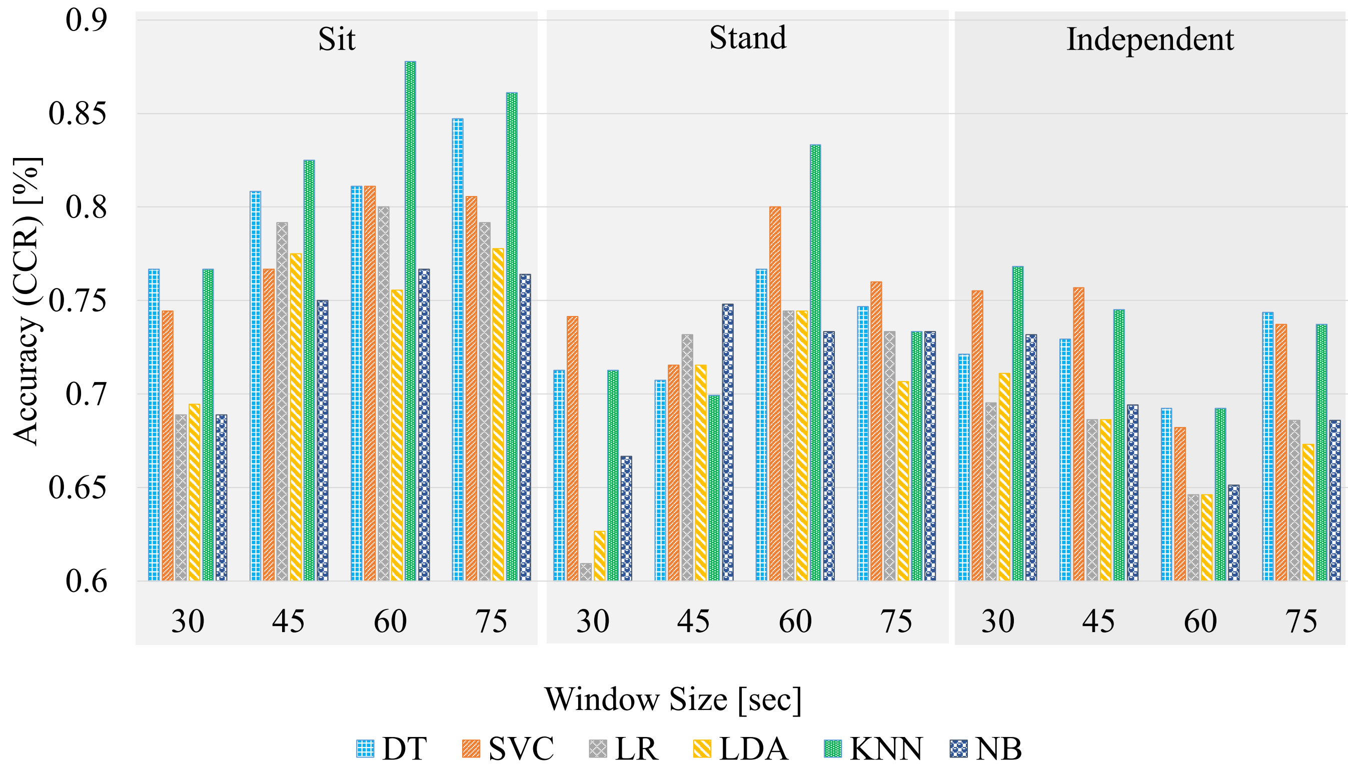


Fig. 3. Performance comparison of different algorithms in different body postures at different window sizes.

presented in Table III, K-NN gives an accuracy of 87.78% for the HL detection in sitting posture scenario. On the other hand, for the standing and posture independent scenarios, the K-NN produces an accuracy of 83.33% and 76.82% for feature combinations D and E, respectively. KNN yields these results for the sets of tuning parameters as follows: 1) for sitting posture, metric: *Minkowski*, $K = 15$, weights: *uniform*; 2) for standing posture, metric: *Manhattan*, $K = 10$, weights: *uniform*; 3) for posture independent scenario, metric: *Manhattan*, $K = 60$, weights: *distance*.

Once the best performing models are found for all three posture based scenarios, then their performance is evaluated against advance evaluation metrics which are presented in Table IV. It can be seen that sitting scenario data shows the highest precision, sensitivity, specificity, and overall accuracy which are 0.81, 0.95, 0.82, 0.88 respectively. From Fig. 4 it can be observed that the best performing K-NN based model, with the parameters mentioned above, exhibits TPR of 0.95 and TNR of 0.82 in sitting posture scenario. It means model could identify 95 % of the dehydrated state instances and 82 % of the hydrated state instances correctly. Similarly, for the standing data, a TPR of 0.92 and a TNR of 0.76 are achieved. Whereas for the data of posture independent scenarios 0.84 TPR and 0.71 TNR are achieved. It can be observed that TPR of dehydrated state detection is even higher than the overall accuracy. Hence, if a reminder or alarm system is set for dehydrated state detection, i.e alarm goes off only when

TABLE IV
PERFORMANCE OF THE BEST PERFORMING K-NN CLASSIFIER AGAINST DIFFERENT METRICS IN EACH POSTURE SPECIF SCENARIO

Posture	Window Size [sec]	Precision	Sensitivity	Specificity	CCR
Sit	60	0.81	0.95	0.82	0.88
Stand	60	0.76	0.93	0.76	0.83
Independent	30	0.71	0.84	0.71	0.77

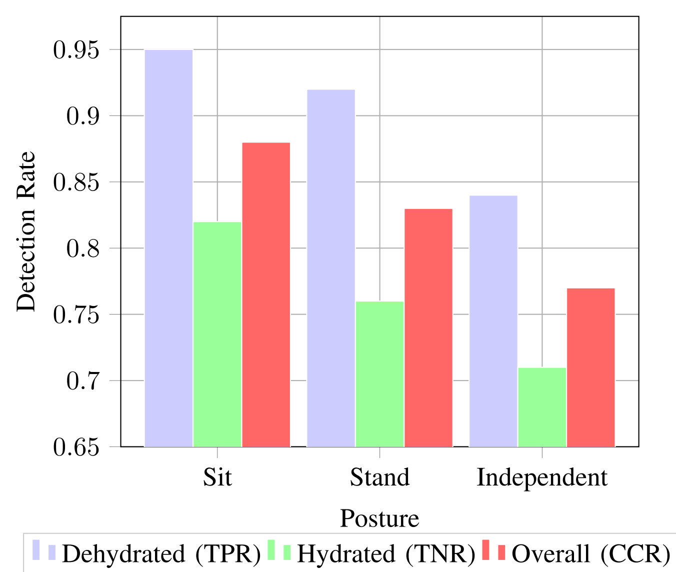


Fig. 4. Class specific performance of the KNN based best models in each posture specific scenario

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

dehydrated state is detected, then less number of false alarm can be expected. It is also observed that algorithms commonly perform better on the window size of 60 seconds for the posture specific scenarios. It is, therefore, safe to say that 60 seconds is an optimal window size for the feature extraction for posture specific scenarios. Another common behaviour observed for almost all algorithms, as can be seen from Fig. 3, is that they generally perform better on posture specific data. It reflects that GSR not only varies because of HL but also because of body posture.

Based on these results, it can be stated that, with the selection of appropriate window size, features set and ML algorithm, GSR data can be used to identify the HL in the human body. More specifically, it can be concluded if GSR data is collected for 60 seconds in sitting posture from an individual, the K-NN model can detect with an accuracy of almost 95% if the individual is dehydrated.

V. Analysis and Discussion

Various factors that directly impact the accuracy of HL estimation models are discussed in this section.

A. *Impact of Postures*

The impact of posture on HL estimation is studied by developing independent models for sitting, standing, and posture independent scenarios. As it can be seen from Fig. 3, the best performing model turns out to be K-NN, for window size on the sitting posture data, with an accuracy of 87%. It can also be seen that overall accuracy for the detection of the HL for the data of sitting posture is between 70% and 88%. It is better than that for the data of the other two scenarios. For the standing and posture independent scenarios overall performance for the most of the algorithms is mediocre, in the range of 60% to 75% except for the few algorithms. Some algorithms like K-NN and SVM marginally perform better for the 60 seconds window on the standing data and 30 seconds window for the independent scenario. No other posture specific trend could be found for standing and independent scenarios. For the both scenarios, performance of the algorithms vary over the window size without any specific pattern. For some algorithms, accuracy is better for standing data, while it is better for posture independent data for other algorithms. All metrics scores are better for sitting posture as compared to the data from the other two scenarios. Comparatively low noise in the sitting scenario can be the reason. The data of standing scenario may have more noise because EDA is very sensitive to internal and external stimuli, like small movements of body muscles.

B. *Impact of Window Size*

Window size, the length of time for which skin conductance data should be measured or considered for feature extraction is an important parameter for model development. It directly impacts the underlying patterns, features and subsequently the performance of the algorithm. To find the best window size, accuracy for the identification of HL is assessed and compared for different window sizes. From Fig. 3 it can be seen that for the posture specific scenarios, use of 60 seconds window produces the best results for the HL estimation with K-NN algorithm. However, for posture independent scenario, window size, 30 seconds exhibits the best accuracy with the K-NN algorithm.

Overall, the performance of the algorithms varies with window sizes for the posture specific and posture independent scenarios. For the sitting posture, performance of the most of algorithms like DT, SVC, LR, K-NN and NB improves with increase in window size from 30 seconds to 60 seconds, except the performance of LDA. For window sizes 60 to 75 seconds, it decreases marginally for all the algorithms, except for DT and LDA which show slightly better performance. Similarly, for most of the algorithms in standing posture scenario, the performance improves with the increase in window size from 30 to 60 seconds with few exceptions for window sizes 60 to 75 seconds where it decreases. For the posture independent scenario, the trend is reversed, the performance of almost all the algorithms seem to degrade from window size 30 to 60 seconds and at window size 75 seconds the performance improves again.

It can be concluded that from window sizes 30 to 60 seconds, for the most of the algorithms, performance improves for posture specific scenario but it decreases for posture independent scenario. Increase in performance with the window size can be correlated with the availability of more data and more information for the extraction of features. The decrease in the performance is an interesting behaviour, found in the posture independent scenario. One potential reason may be the confusion of posture specific characteristics with the state-specific characteristics. Posture independent data includes the data sets for both scenarios (i.e. sitting and standing) and GSR is found to vary with the change in posture as well. So posture specific variations may be confused with the variations related to HL, which in turn reduces the overall accuracy for the posture independent scenario.

C. *Impact of Feature Combinations*

Identification of valuable features and selection of the right combination of features is also a crucial task that highly impacts the performance of the models. The main reason of using statistical features is that they are easy to extract and use in the ML models. A simple feature extraction process can reduce the computational cost and time of processing, which is also important for real-time models.

The performance of each individual feature is evaluated for all the posture based scenarios, window sizes and algorithms. It is observed that few of the features, like standard deviation, variance, entropy, and kurtosis do not perform very well in almost all scenarios and result in an accuracy of around 50% individually. Other features, like minimum, mean, median, mode, and percentile, individually perform better nearly in all situations with an overall average accuracy of around 66%. Maximum accuracy achieved by a single individual feature is around 82% and it is for the feature ‘mean’ and algorithm K-NN applied with the window size of 75 seconds on the sitting data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

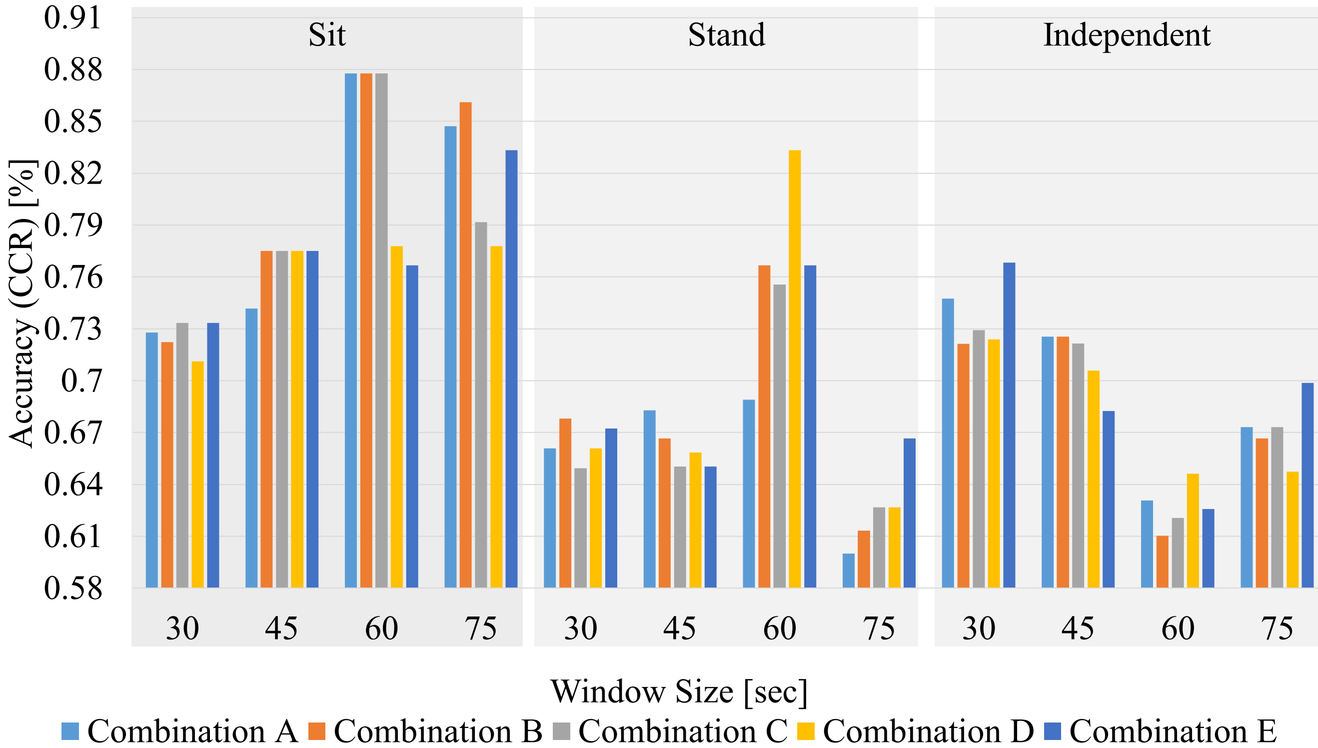


Fig. 5. Comparison of the performance of K-NN for a different combination of features in different body postures and different window sizes. Please refer to Table III for the details about the combinations.

However, the use of a single feature is not good enough to achieve the best possible accuracy. To improve the accuracy combination of features can be used. All possible combinations of features are also evaluated for all the possible situations. It is found that performance varies with different sets of features for different schemes of data, window sizes and algorithms. although such combinations can improve the performance significantly for some individual scenarios, but no single combination could outperform all other combinations for all scenarios. The top-scoring five feature combinations are presented in Table III along with the best accuracy they achieve against the specific window size and posture. Performance comparison of those sets, which yield the highest accuracy, is shown in Fig. 5. The comparison is presented for their performance with the K-NN algorithm for different window sizes and posture based scenarios. Only the performance for K-NN is presented here because it is the algorithm which gives the best accuracy overall.

Interesting findings from the feature analysis also include that the individually better-performing features not necessarily perform well when they are tried in combination with other good features, on the contrary, low performing features along other features can improve the performance. For example, 'kurtosis' individually gives an accuracy between 42% and 59%, but when it is used with in combination with other features the performance of the algorithms can improve. It is also observed that a feature combination which generates the best performance for one scenario does not exhibit impressive performance for the other scenarios. One possible underlying reason can be the posture specific variations in the data. For example, the standing data has more variations and pikes due to small movements in hydrated state, but it is not the case for dehydrated state in standing posture. Besides that, data for sitting posture is steady and stable overall for the both HL as compared to standing posture data. This inference is supported by the presence of 'kurtosis' feature, a measure for peaks, in the best combination D for standing data. Whereas, all top-scoring combinations for sitting data do not use this feature.

From Table III it can be seen that top scorer model for sitting posture has four common features. Out of those four, 'mean', 'entropy', and 'standard deviation' are also present in the best model for standing posture. However, the standing posture model has some additional features that may represent the standing posture specific variations related to HL. The best posture independent model has two features common with sitting and standing models, another feature just common with standing model and one feature of its own, 'variance'. The 'minimum' feature, common with the best standing posture model, can be a cause of low performance for the independent data scenario because HL and posture specific characteristics may be confused together.

D. Impact of Machine Learning Algorithms

Six popular classification algorithms of ML are implemented and evaluated for all posture based scenarios, window sizes and feature combinations to find the best performing algorithm. The main objective has been to identify a single algorithm which may perform better for all scenarios. As

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

reflected in Fig. 3, it is found that K-NN is the only algorithm which outperforms other algorithms in the posture specific as well as posture independent scenarios. It is also observed that the performance of the K-NN at its best is far superior to other algorithms particularly on sitting and standing data. For example, for sitting data and standing data, at window size of 60 seconds, the next ten best performing models are also K-NN based models with different combinations of features. Whereas for the posture independent scenario K-NN performs marginally better, at window size of 30 seconds, but the overall performance is low here, for almost all the algorithms as compared to other postures.

Almost all algorithms perform better on sitting data, it can be associated with the better quality of data rather than the performance of algorithms. Except for K-NN, there are two algorithms SVC and DT that consistently perform better across almost all window sizes and for all posture based scenarios. NB, the probabilistic algorithm, and LR and LDA, linear classifiers, could not perform very well. One possible reason is that the data points do not have very clear linear boundaries for HL whereas these algorithms produce linear boundaries for the segregation of classes. On the other hand, K-NN and DT could perform better because instead of defining linear boundaries they can form more complex boundaries or planes for separating hydrated and dehydrated cases. They group together data items for their similarity with other data points. Similarly, SVC with its range of cost functions, including linear and non-linear, is also a robust classifier and can be the second choice here. If the model is required to perform in real-time, SVC can be considered for an advantage it has over the K-NN. It is more efficient for the real-time processing because SVC uses only the trained model whereas the K-NN also needs data samples in memory to compare with the new data points which make it inefficient for the real-time processing particularly for low energy devices with limited computation and storage. However, in the case of SVC, a compromise is to be made on the overall accuracy which is slightly low as compared to K-NN.

VI. Conclusions

HL monitoring is very important for maintaining an appropriate HL in human body. There are very limited solutions available for the estimation of HL which also have drawbacks like being invasive, inconsistent, inaccurate and unreliable. In this paper, a non-invasive solution is proposed for the auto-detection of the HL in the human body. GSR data is collected in different scenarios like sitting and standing. Statistical features are extracted from GSR data, over four different windows of time for each posture based scenario. Six popular classification algorithms are implemented on various combinations of the features, to find the optimal window size, feature combination and algorithm which give the best accuracy for HL detection in the human body. It is found that K-NN outperforms all other algorithms. It gives a CCR score of 87.78% and 83.33% for the sitting and standing postures respectively, for the window size of 60 seconds. For a posture independent scenario CCR is 76.82% for a window size of 30 seconds. True positive rate are even higher like 0.95, 0.92 and 0.84.71 for sitting, standing and posture independent scenarios respectively. It is concluded that for GSR data collected for 60 seconds in sitting posture, the K-NN based model can detect with an accuracy of almost 95% if an individual is dehydrated. In future this work can be extended to study the hydration level detection as multi-class problem instead of binary class. For that purpose advance ML algorithms like deep learning will be considered.

Acknowledgment

This work is funded by project AARE17-019 provided by the ADEC Award for Research Excellence, Abu Dhabi, United Arab Emirates University.

References

[1] Ahmed M El-Sharkawy, Opinder Sahota, and Dileep N Lobo. Acute and chronic effects of hydration status on health. *Nutrition reviews*, 73(suppl_2):97–109, 2015.
[2] David C Garrett, Nyssa Rae, Jared R Fletcher, Sasha Zarnke, Sarah Thorson, David B Hogan, and Elise C Fear. Engineering approaches to assessing hydration status. *IEEE reviews in biomedical engineering*, 11:233–248, 2017.
[3] Sunayana S Kakade and Amol B Jagadale. Development of system for estimation of total body water (tbw), fat mass (fm), fat free mass (ffm) using bioimpedance analysis technique. In *2016 International Conference on Communication and Signal Processing (ICCSP)*, pages 0224–0226. IEEE, 2016.
[4] Raymond M Fish and Leslie A Geddes. Conduction of electrical current to and through the human body: a review. *Eplasty*, 9, 2009.
[5] Leon Thomas Lee Marsh. Hydration monitor, November 5 2013. US Patent 8,574,165.
[6] Gerald Sweeney, Cory McCluskey, and James William Pfeiffer. Valve and cap system for a beverage container, December 9 2014. US Patent 8,907,796.
[7] Amanda Myers, Abhishek Malhotra, John Muth, Yong Zhu, and Shan-shan Yao. Personal hydration monitor, November 24 2016. US Patent App. 15/160,594.
[8] A Rizwan, A Zoha, A Alomainy, NA Ali, MA Imran, and QH Abbasi. Skin conductance as proxy for the identification of hydration level in human body. In *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, volume 1, pages 1–3. IEEE, 2019.
[9] Jason J Braithwaite, Derrick G Watson, Robert Jones, and Mickey Rowe. A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments. *Psychophysiology*, 49(1):1017–1034, 2013.
[10] Xian Huang, Huanyu Cheng, Kaile Chen, Yilin Zhang, Yihui Zhang, Yuhao Liu, Chenqi Zhu, Shao-chi Ouyang, Gil-Woo Kong, Cunjiang Yu, et al. Epidermal impedance sensing sheets for precision hydration assessment and spatial mapping. *IEEE Transactions on Biomedical Engineering*, 60(10):2848–2857, 2013.
[11] Hugo Plácido Da Silva, José Guerreiro, André Lourenço, Ana LN Fred, and Raúl Martins. Bitalino: A novel hardware framework for physiological computing. In *PhyCS*, pages 246–253, 2014.
[12] Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, Wolfram Boucsein, Don C Fowles, Sverre Grimnes, Gershon Ben-Shakhar, Walton T Roth, Michael E Dawson, and Diane L Filion. Publication recommendations for electrodermal measurements. *Psychophysiology*, 49(8):1017–1034, 2012.
[13] Félix Garcıa López, Miguel Garcıa Torres, Belén Melián Batista, José A Moreno Pérez, and J Marcos Moreno-Vega. Solving feature subset selection problem by a parallel scatter search. *European Journal of Operational Research*, 169(2):477–489, 2006.
[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
[15] Mohammad Hossin and MN Sulaiman. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1, 2015.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

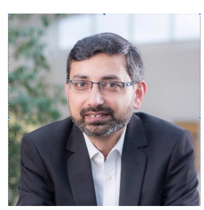
ALI RIZWAN received the bachelor's degree in applied and theoretical math in 2006 and the MBA-IT degree from Bahauddin Zakariya University, Pakistan, in 2008, and the M.Sc. degree in big data science from the Queen Mary University of London, London, U.K., in 2016. He is currently pursuing the Ph.D. degree with the University of Glasgow, Glasgow, U.K. He is conducting research, in the area of big data analytics for self-organizing wireless networks. His other area of interest is big data analytics for healthcare.

Najah Abu Ali is currently a Professor at the Faculty of Information Technology in the United Arab Emirates University (UAEU). She received her Ph.D. from the Department of Electrical and Computer Engineering at Queen's University in Kingston, Canada; specializing in resource management in computer networks. Her MSc and BSc were both attained in Electrical Engineering at the University of Jordan. Her general research interest includes modeling wireless communications, resource management in wired and wireless networks, and reducing the energy requirements in wireless sensor networks. More recently, she has strengthened her focus on the Internet of Things, particularly at the nano-scale communications, in addition, vehicle-to-vehicle networking. Her work has been consistently published in key publications venues for journals and conference. She's further co-authored a Wiley book on 4G and beyond cellular communication networks. She has also delivered various seminar and tutorials at flagship IEEE Communication Societies. She has also served as a reviewer for many noted journals, a TPC member and reviewer in several conferences (including ICC and Globecom), and contributed to the organization of several SIs and workshops. She also served as the Publication co-chair of IEEE GC'18.

Ahmed Zoha is a lecturer at the James Watt School of Engineering, University of Glasgow, UK,and research investigator with Scotland 5G Center. He received his PhD from the Institute of Communication Systems (ICS) at the University of Surrey and MSc in Communication Engineering from Chalmers University of Technology, Sweden. Dr Ahmed has more than 10 years of experience in the domain of quantitative analysis and research, data mining, predictive modelling, and applied machine learning and artificial intelligence for real-world applications. Dr Ahmed has contributed to several interdisciplinary and multimillion-funded international research projects in the domain of big data enabled self-organizing wireless communication networks, smart energy monitoring and connected healthcare. His research interests include machine learning at the edge, 5G and beyond, digital health and internet of everything.

Metin Ozturk received the B.Sc. degree in Electrical and Electronics Engineering from Eskisehir Osmangazi University, Turkey, in 2013, and the M.Sc. degree in Electronics and Communication Engineering from Ankara Yıldırım Beyazıt University, Turkey, in 2016. He also worked as a research assistant at Ankara Yıldırım Beyazıt University from 2013 to 2016. He is currently pursuing his Ph.D. degree with the James Watt School of Engineering, University of Glasgow, United Kingdom. His research interests include Self-Organizing Networks (SON), mobile communications, and machine learning algorithms.

AKRAM ALOMAINY received the M.Eng. degree in communication engineering and the Ph.D. degree in electrical and electronic engineering (specialized in antennas and radio propagation) from Queen Mary University of London (QMUL), U.K., in July 2003 and July 2007, respectively. He joined the School of Electronic Engineering and Computer Science, QMUL, in 2007, where he is a Reader in Antennas Applied EM. His current research interests include small and compact antennas for wireless body area networks, radio propagation characterisation and modelling, antenna interactions with human body, computational electromagnetic, advanced antenna enhancement techniques for mobile and personal wireless communications, nano-scale networks and communications, THz material characterisation and communication links and advanced algorithm for smart and intelligent antenna and cognitive radio system. He has authored and co-authored four books, 6 book chapters and more than 350 technical papers (6400+ citations and H-index 35) in leading journals and peer-reviewed conferences. Dr Alomainy won the Isambard Brunel Kingdom Award, in 2011, for being an outstanding young science and engineering communicator. He was selected to deliver a TEDx talk about the science of electromagnetic and also participated in many public engagement initiatives and festivals. He is an elected member of UK URSI (International Union of Radio Science) panel to represent the UK interests of URSI Commission B (1 Sept 2014 until 31 Aug 2020).



MUHAMMAD ALI IMRAN (M'03–SM'12) received the M.Sc. (Hons.) and Ph.D. degrees from Imperial College London, U.K., in 2002 and 2007, respectively. He is Dean Glasgow College UESTC and a Professor of communication systems with the James Watt School of Engineering, University of Glasgow, U.K. He is an Affiliate Professor at the University of Oklahoma, USA, and a Visiting Professor at the 5G Innovation Centre, University of Surrey, U.K. He is leading research in University of Glasgow for Scotland 5G Center. He has over 18 years of combined academic and industry experience, working primarily in the research areas of cellular communication systems. He has been awarded 15 patents,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.2965892, IEEE Sensors Journal

has authored/co-authored over 400 journal and conference publications, and has been principal/co-principal investigator on over £6 million in sponsored research grants and contracts. He has supervised 40+ successful Ph.D. graduates. He has an award of excellence in recognition of his academic achievements, conferred by the President of Pakistan. He was also awarded the IEEE Comsoc's Fred Ellersick Award 2014, the FEPS Learning and Teaching Award 2014, and the Sentinel of Science Award 2016. He was twice nominated for the Tony Jean's Inspirational Teaching Award. He is a shortlisted finalist for The Wharton-QS Stars Awards 2014, the QS Stars Reimagine Education Award 2016 for innovative teaching, and VC's Learning and Teaching Award from the University of Surrey. He is a Senior Fellow of the Higher Education Academy, U.K. He is the editor/co-editor of 8 books.

QAMMER H. ABBASI (SM'16) received the B.Sc. and M.Sc. degrees in electronics and telecommunication engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, and the Ph.D. degree in electronic and electrical engineering from the Queen Mary University of London (QMUL), U.K., in 2012. In 2012, he was a Postdoctoral Research Assistant with the Antenna and Electromagnetics Group, QMUL. He is currently a Lecturer (Assistant Professor) with the James Watt School of Engineering, University of Glasgow, U.K. and researchinvestigator with Scotland 5G Center. He has contributed to over 250 leading international technical journal and peer reviewed conference papers, and eight books. He received several recognitions for his research, which include appearance on BBC, STV, dawnnews, local and international newspapers, cover of MDPI journal, most downloaded articles, U.K. exceptional talent endorsement by Royal Academy of Engineering, National Talent Pool Award by Pakistan, International Young Scientist Award by NSFC China, URSI Young Scientist Award, National Interest Waiver by USA, four best paper awards, and best representative image of an outcome by QNRF. He is an Associate Editor for the IEEE JOURNAL OF ELECTROMAGNETICS, RF AND MICROWAVES IN MEDICINE AND BIOLOGY, the IEEE SENSORS JOURNAL, IEEE OPEN ACCESS ANTENNA AND PROPAGATION, IEEE ACCESS and acted as a guest editor for numerous special issues in top notch journals.